Aalto University

School of Science

Degree Programme in Industrial Engineering and Management

Olli Kaplas

Airport capacity modeling

Case Helsinki Airport

Master's Thesis

Espoo, May 2014

Supervisor: Eero Eloranta, Prof.

Instructor: Taija Turunen, D.Sc.(Tech)

Aalto University School of Science

Aalto University School of Science Degree Programme in Industrial Engineering and Management		ABSTRACT OF THE MASTER'S THESIS
Author: Olli Kaplas		
Title: Airport capacity modeling – Case Helsinki Airport		
Number of pages: 86 (95)	Date: 13.5.2014	Language: English
Professorship: Industrial Management		Code: TU-22
Supervisor: Eero Eloranta, Prof.		
Instructor(s): Taija Turunen, D.Sc. (Tech)		

Abstract:

This study focuses on the modeling of airport processes at Helsinki Airport in the perspective of capacity and utilization. Objective of the research is to create a model of the airport as a system of processes and model the capacities of selected individual processes. The possibility of emerging bottleneck in the selected processes is assessed. Conceptual modeling is used to depict the complete airport in operational point of view. In the modeling of runway capacity a method based on statistical method of quantile regression is used. In the deicing capacity modeling a gamma regression model of the generalized linear models is used. In the baggage processing capacity general statistical analysis is used.

A process diagram model is created of the main processes as a system at Helsinki Airport. The model includes the main flows of objects at an airport (aircraft, baggage and passengers) from processes or states to other processes and states. The diagram helps to clarify on a high level how the relationships of how capacities of the consecutive process are interlinked.

The capacity of the runway process is modeled and the effects of the used runway configuration, snowfall, wind speed and visual conditions are assessed in terms of the runway capacity at Helsinki Airport. The runway capacity seems to be sufficient in most of the cases except when some weather factors are affecting it. The runway capacity seems to be a bottleneck for the operation in heavy snowfall conditions.

The deicing treatment capacity is studied by creating a model explaining and forecasting the deicing treatment time. On the basis of the deicing time model an equation for calculating expected deicing capacity as a function of the modeled parameters is formulated. The model for deicing capacity would suggest that in most cases 10 trucks would suffice the current demand and avoid the deicing treatment of becoming a bottleneck.

The baggage process at Helsinki Airport was studied and two points in the process were identified as constraining for the whole process. One of them, the transfer baggage loading lines, is studied more deeply by studying the distribution of the load on the capacity. The analysis of the baggage capacity would suggest that around an increase of 600 baggage/hour could be possible before its capacity reaches its limits at Helsinki Airport and starts to be a bottleneck for the operation.

Keywords: capacity, bottleneck, airport, modeling, business processes, systems thinking, theory of constraints, service supply chain network, quantile regression, gamma regression, generalized linear models

Aalto-yliopisto Perustieteiden korkeakoulu Tuotantotalouden koulutusohjelma	DIPLOMITYÖN TIIVISTELMÄ	
Tekijä: Olli Kaplas		
Työn nimi: Lentokentän kapasiteetin mallintaminen – Case Helsinki-Vantaan lentoasema		
Sivumäärä: 86 (95)	Päiväys: 13.5.2014	Julkaisukieli: Englanti
Professuuri: Teollisuustalous	Professuurikoodi: TU-22	
Työn valvoja: Eero Eloranta, Prof.		
Työn ohjaaja(t): Taija Turunen, TkT		

Tiivistelmä:

Tämä diplomityö keskittyy lentokentän prosessien mallintamiseen Helsinki-Vantaan lentokentällä kapasiteetin ja käyttöasteiden näkökulmasta. Tutkimuksen tavoitteena on lentokentän mallintaminen sen pääprosessien muodostamana systeeminä sekä valittujen prosessien kapasiteettien mallintaminen. Pullonkaulojen Konseptuaalista mallintamista käytetään kuvaamaan koko lentokenttää operatiivisesta näkökulmasta. Kiitoratojen kapasiteetin mallintamiseen käytetään tilstollista menetelmää, joka perustuu kvantiiliregressioon. Deicing-käsittelyn kapasiteetin mallintamisessa käytetään tilastollisten yleistettyjen lineaaristen mallien joukkoon kuuluvaa gamma regressiota. Matkatavaroiden prosessoinnin kapasiteettia mallinnetaan yleisellä tilastollisella analysoinnilla.

Helsinki-Vantaan lentoasemasta luotu prosessikaaviomalli kuvaa lentokenttää pääprosessien muodostamana systeeminä. Malli kuvaa ensisijaisten objektien virtaukset (lentokoneet, matkatavarat ja matkustajat) lentokenttäjärjestelmän sisällä paikasta ja tilasta toiseen. Kaaviomalli auttaa selvittämään miten perättäisten prosessien kapasiteetit ovat linkittyneet toisiinsa.

Kiitorataprosessin kapasitetti mallinnetaan ja käytössäolevan kiitoratakonfiguraation, sataneen lumen, keskimääräisen tuulennopeuden ja näkyvyyden vaikutukset kapasiteettin arvioitiin mallinnuksen avulla. Kiitoratojenkapasiteetti näyttää olevan lähes riittävä suurimmassa osassa tapauksia lukuunottamatta tilanteita joissa tietyt sääolosuhteet vaikuttavat kapasiteettia alentavasti.

Deicing-käsittelyn kapasiteettia tutkitaan muodostamalla malli deicing-käsittelyn kestolle. Deicicing-käsittelyn keston mallin avulla työssä luodaan kaava odotettavissa olevan kapasiteetin laskemiselle parametrien arvojen suhteen. Deicing-käsittelyn kapasiteetti viittaisi siihen, että lähes aina 10 samanaikaista deicing-käsittelyrekkaa riittää täyttämään lentokentän nykyisen lähtevien lentojen deicing-käsittelytarpeen, ennen kuin prosessista muodostuu pullonkaula.

Matkatavarankäsittelyprosessia kapasiteettia tutkitaan ja kaksi kohtaa prosessissa tunnistetaan mahdollisiksi pullonkauloiksi. Toista niistä, siirtomatkatavaroiden lastauslinjoja, selvitetään tarkemmin tutkimalla kapasiteetin kuormaa kyseisessä kohdassa. Analyysi viittaisi siihen, että tässä kohdassa keskimääri n. 600 laukkua/tunti lisäys olisi mahdollinen, ennenkuin kapasiteetti saavuttaa ylärajansa matkatavaroiden käsittelyssä ja muodostaa pullonkaulan operaatioille.

Asiasanat: kapasiteetti, pullonkaula, lentokenttä, mallintaminen, liiketoimintaprosessit, systeemiajattelu, kapeikkoajattelu, palvelutoimitusverkosto, kvantiiliregressio, gamma regressio, yleistetyt lineaariset mallit

Preface

This thesis was done as a part of research effort of research program Future Industrial Services (*FUTIS*) in Service Engineering and Management (SEM) research group at BIT-Research Centre of the Department of Industrial Engineering and Management at Aalto University. The purpose of this master's thesis is to study and model the capacity at Helsinki Airport. Helsinki Airport constitutes a service supply chain network where different service suppliers affect the overall capacity of the airport. The capacity at Helsinki Airport was modeled as whole and three selected processes were modeled and analyzed separately with statistical methods. This thesis introduces the reader how the capacities were modeled, what the results are, what kinds of assumptions were used, what kind of data was used and what the possible error sources are.

During the course of creation of this thesis I've learned a lot about aviation business which was new to me in other ways than as a customer. I also got to see how the Helsinki airport operates behind the scenes, which was very interesting. I've found myself developing during this whole process and learned whole new ways of thinking. The process of conducting this study was a remarkable journey.

I want to thank Finnair and especially Ville Iho for coming up with the idea of studying the capacity at Helsinki Airport and giving the assignment for this thesis. I also want to thank my instructor at Finnair, Janne Tarvainen, for instructing the thesis in practical and operational point of view. I want to thank Jukka Glader for arranging thorough introduction to the operation at Helsinki Airport. I would like to thank Finnair's operational analysts Hanna Salmi and Juha Karstunen for brainstorming ideas and for giving input to the contents of this thesis and for helping in data gathering.

I would like to thank my instructor Taija Turunen and my supervising Professor Eero Eloranta for giving scientific counseling and support during the course of writing this thesis.

I want to thank my future mother-in-law Anna-Kaarina Kairamo for proofreading the text I produced. Lastly I want to thank Milja, my fiancée, for supporting me in the process of writing this thesis.

Terminology and abbreviations

The terms and abbreviations used in this study are defined and introduced here.

Apron = The area at an airport where the aircraft are parked, unloaded, loaded, fueled and boarded.

AC = Aircraft

ATC = Air Traffic Control

BLC = Baggage Logistics Centre at Helsinki Airport

***EFHK* =** A unique code name set by *ICAO* for the Helsinki-Vantaa airport to distinguish it from all other airports in the world.

EASA = European Aviation Safety Agency is an organization, who has the legal authority in EU to make regulations concerning the safety of aviation.

ICAO = International Civil Aviation Organization is a specialized agency of the United Nations to promote the safe and orderly development of international civil aviation throughout the world. It sets standards and regulations necessary for aviation safety, security, efficiency and regularity, as well as for aviation environmental protection.

***IFR* =** Instrument flight rules govern flight under conditions in which flight by outside visual reference is not safe.

***VFR* =** Visual flight rules are a set of regulations under which a pilot operates an aircraft in weather conditions generally clear enough to allow the pilot to see where the aircraft is going, as specified in the rules of the relevant aviation authority. The pilot must be able to operate the aircraft with visual reference to the ground, and by visually avoiding obstructions and other aircraft.

Runway operation = Either an arrival or a departure from a runway.

TOC = Theory of Constraints, a theory on the constraining resources of a system and their implications on the overall capacity of system introduced by Goldratt and Cox in their book the Goal (Goldratt and Cox, 1993).

Contents

List of figures

List of tables

1. Introduction

Finavia, the airport operator of Helsinki Airport, has announced plans to extend the airport with large investments and it will have great impact on the capacity in the future. Indeed the passenger amounts at Helsinki Airport have increased from 9,7 million in 2003 to 15,3 million in 2013 (Finavia Oyj, 2013a). Given the daily cyclical demand of flights at Helsinki airport, the momentary strain on the capacity can be even more than this figure would suggest. Finnish airline Finnair has projected that the Asian flight revenues will double by between 2010 and 2020 (Finnair Oyj, 2013), which puts requirements for even more capacity from the airport in the future.

New business models in the airline business have started to emerge from the 1990s onwards (Doganis, 2001,p.213). Many airline companies have been moving away from the traditional airline model, where airlines have separate functions or departments to provide all the services they need (Doganis, 2001, p.214). There has been a growing trend in the airline business to changing of the business models and focusing on the core competences of the airlines to reflect this. It has meant outsourcing of many of the functions that have traditionally been provided by the airline companies. This trend has been seen in Helsinki Airport where Finnair, the airline with the most operations, has also changed its business model in the course of time. At Helsinki Airport this has meant that the service supply chain network has grown and become more complex in the perspective of the organizations involved.

This study was started by the interest of Finnair on the efficient operation of Helsinki airport. When I first started working on this thesis the interest was to study the capacities and processes at Helsinki airport, model them and try to figure out when the processes become bottlenecks for the system in whole. The idea was that the airport could be modeled like a factory. The airport can be seen as a system where a bottleneck is a process that is constraining the operation of the system in a certain state of the system. Processes have certain limits in their performance, namely their capacities are limited, and the limits are many times dependent on multiple factors in the airport environment. When the demand exceeds the capacity threshold for a certain process, the process can become a bottleneck. The capacities of the different processes have a major effect on how well the system plays as a whole.

A system is a group of interacting, interrelated components that form a complex and unified whole (Anderson and Johnson, 1997). An airport certainly is a complex system with a lot of components. The systems components can be physical objects or they can be intangible such as processes, relationships, company policies, information flows, interpersonal interactions and internal state of mind such as feelings, values and beliefs (Anderson and Johnson, 1997). An airport includes all of these types of components, and thinking of Helsinki airport, it is a wonder how the complex system manages to operate so successfully from day to day.

The fundamental mission of an airport as a system is getting the passengers, baggage, freight and aircraft to the right place at a right time with best possible end customer experience. To achieve this, an airport incorporates large service supply chain of many organizations and stakeholders operating and influencing the operation at the airport. As such the operation of an airport is also affected by outside influences like the operations on other airports and environmental influences like meteorological conditions.

An airport as system can be seen as having three flows of objects of primal interest. Successful control and management of these flows in the system defines the success of the fundamental mission or purpose of an airport. The main flows of objects are the flows of passengers, flows of baggage and freight and the flow of aircraft. The key in successful operation of the airport is to manage what processes are performed, how the processes are performed, how the timing and duration of the processes' activities are managed and in what order activities are performed and where the processes take place.

In this study a comprehensive model of Helsinki airport as a system illustrating the main flows of objects and processes affecting these was created. On the basis of the model, one can deduct the relationships and interdependencies between consecutive processes in the perspective of the capacity. Also concerning the capacities of processes this study focuses on three processes in more detail: the runway process, baggage handling facility (Baggage Logistics Centre) and the deicing process which takes place in the winter.

1.1 Research background – The operational efficiency of an airport

This research was initiated by the Finnish airline company Finnair. The need for this kind of study has come due to the growing passenger amounts at the Helsinki-Vantaa airport and need to study in more detail the capacities of the processes at airport and how they are affected by different conditions. The grown passenger amounts have resulted in longer flight delays, which is an indication of a possible outreach of capacity at an airport.

Airport constitutes a complex service system. There are lots of organizations providing services and the customer-supplier relationships are multiple. Most important organizations in the airport service supply chain at Helsinki Airport are depicted in Figure 1. Organizations affecting the regulations under which organization operate at an airport, are depicted with red color and organizations conducting and affecting the operations directly at the airport are depicted with blue color.

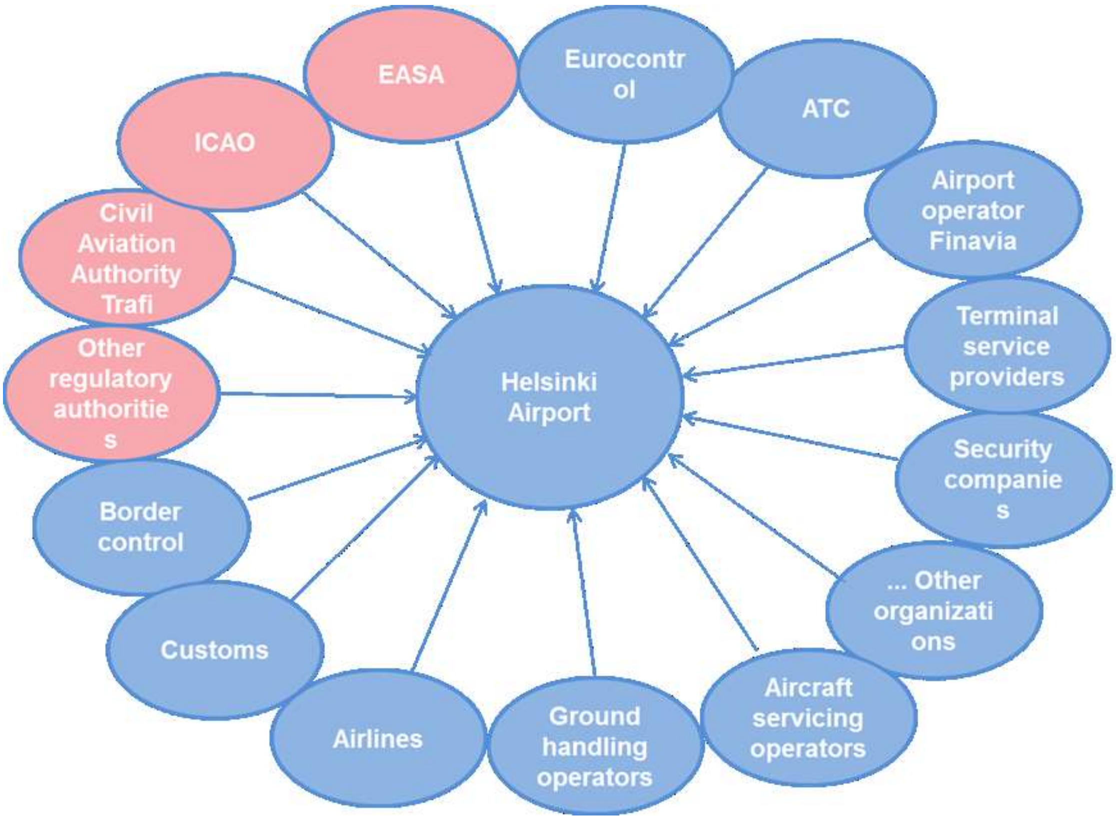


Figure 1: Organizations influencing and acting on the service system of Helsinki Airport

The service suppliers operating at an airport constitute a service supply network, where different organizations provide different services. An illustration of the main customer-supplier relationships in the regulatory environment of aviation ending to the services provided to the end customer is depicted in Figure 2. The aviation business is heavily regulated and it deeply affects the service supply chain of Helsinki airport. It is why the regulatory environment set by the regulatory authorities is depicted.

The airport operator Finavia is the most influential organization when considering the efficiency of operation at the airport. It is of course natural since they are the airport operating (and owning) company. The airport operator is in multiple customer-supplier relationships in the service supply network, because of its central role at the airport.

The air traffic control (*ATC*) is an organization within the airport operating company Finavia, but it is listed separately because of its major importance at Helsinki Airport. They control and supervise the minute to minute flight operations at the airport and the airspace. Eurocontrol is a European organization providing services and information systems for the airports and airlines operating in Europe. Eurocontrol coordinates and plans and the management or air traffic control for all of Europe.

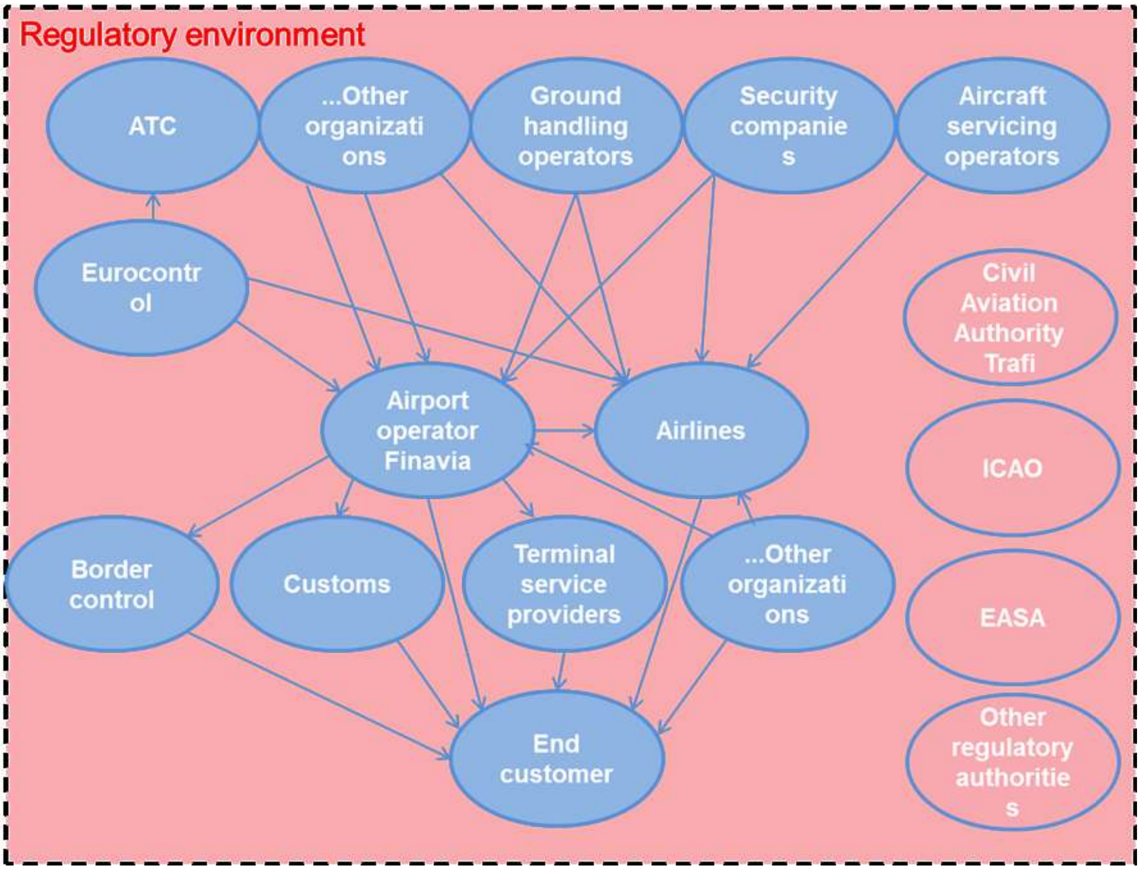


Figure 2: The service supply chain network of main supplier-customer relationship between organizations in the regulatory environment of aviation

The Finnish civil aviation regulatory authority, Trafi, is the national aviation regulatory authority responsible for the regulations of aviation. Trafi's regulations are affected by agreements with other regulation authorities such as the International Civil Aviation Authority (*ICAO*). There's also the European Aviation Safety Agency (*EASA*), who has the legal authority in EU to make regulations concerning the safety of aviation. The regulations and the operations at Helsinki airport are also affected by other regulatory authorities like laws and other regulations set by the government. Together the regulatory authorities have major impact on how the operations at an airport can be conducted and what the service supply chain network of Helsinki Airport is like.

The second most influential organizations affecting the airport efficiency are the airlines. The airlines provide their services to the end customers, which can be consumers or companies. Ground handling operator companies are responsible for the ground services on the aircraft when they are parked. They are service providers for the airlines. The ground handling companies are in key role in everyday operations as their actions can directly induce large delays for the flights and thus affecting the overall efficiency of the operations. The aircraft servicing companies here are the companies like technical, catering and cleaning companies that are servicing the aircraft but are strictly speaking not ground handling companies. The aircraft servicing companies are also suppliers to the airlines.

The customs and border control provide services that are required by the law to the end customer. The end customer has the obligation by law to accept these services. The airport operator Finavia provides facilities for the border control and customs so in a sense they are Finavia's customers. Particularly border control needs to be efficient in operations in order not to disturb the primary operations at an airport. As the customs is the last process before a passenger exits the controlled area at an airport, it is not as critical a process for the whole system.

The airport's terminal service providers are all the service providers that operate inside the terminal and provide services to the passengers. All the shops, bars, restaurants, cafes and such are providing services to the end customer and are airport operating company's customers, renting the spaces from Finavia. Also key organizations affecting the customer experience and the course of the passenger through the terminal are the security companies working throughout the airport. Naturally, their actions affect the efficiency of an airport.

The sheer number of different organizations and their relationship with each other compose major factor affecting the efficiency of the airport operations. Efficiency of the single processes ends up defining how well the mission of the airport is completed. Particularly the capacities of the primary processes have an effect on how well the system plays as a whole.

The capacity can be exceeded in almost any of the single processes conducted by airport's service system, which are directly linked to the airlines provision of air travel service. For instance disruptions in e.g. runway operations, fuelling, baggage loading or security check services might lead to delays for departing flights or bad customer experience. The flight delays can lead to situations where end-customers miss their connecting flights creating costs and possibly resulting in lost revenue when customer doesn't come back. This is why it is of primary interest to the airlines that the operations at the airport are efficient and that the capacities of different processes are sufficient in all given conditions. Bad operation at an airport can draw customers away to use alternative routes. Many airlines have a home airport where majority of their flight legs depart or arrive, which is one of the reasons why Finnair, whose home airport Helsinki airport is, was the initiator of this study and not the airport operator company Finavia.

1.2 Research objectives

It seems that there are no empirical studies made in the literature on how to model a complete airport. An objective of this study is to create model of Helsinki airport as a whole system including the primary processes at the airport. The purpose of the model is to create an understanding of the linkages between different processes and the entities going through them in order to help in assessing the effects of different capacities on the overall system in a factory like manner.

Another objective is to try to figure out and model the capacities of selected processes for the estimation of the capacity measures modeling. As the capacities of operations at an airport fluctuate heavily and are many times subject to different complex endogenous and exogenous effects of different variables, identifying and trying to quantify some of the factors affecting the capacities is set as an objective in this study. The objective is also to try to evaluate the effects of single processes operating at capacity from the whole systems perspective. The processes selected to be in this study are:

- Runway process (the arrivals and departures of aircraft)
- Deicing treatment process
- Baggage process

Finally, this study evaluates under which circumstances the emergence of a bottleneck in the selected processes might take place. When demand for capacity in a certain process is larger than the capacity in the current state of the system, the process might become a bottleneck for the system. The emergence is evaluated under the identified capacities and the identified factors' effects. In these cases the bottleneck is constraining the overall throughput of the system.

1.3 Research questions and research scope

The research questions have been constructed on the basis of the research objectives. They are:

- *How can the operation at Helsinki airport be depicted as a process model?*
- *What are the capacities of the selected processes and how can they be modeled?*
- *What factors affect the capacities of selected processes and what is the size of their impact?*
- *When do the selected processes become bottlenecks for the operation as a whole?*

The airport operation is to be modeled in a high level process model. The purpose of the high level process model is to analyze which of the processes' capacities effect the forming of bottlenecks at Helsinki Airport and in which way they are linked to the overall process. The bottleneck in general sense means that a single process acts as a constraint relative to the other processes. Determining a bottleneck is a complex task when considering the processes even in a high level diagram because of the open system nature of an airport and the complexity of the interlinked processes.

The processes selected to be studied in more detail in this thesis are the runway process capacity, deicing process capacity and the Baggage Logistics Centre (*BLC*) process capacity. The runway process, which consists of arrival and departure operations, is the single most important operation of an airport. This is why it was included in the study as the first process to be studied. The deicing process was selected because of the general notion, that it induces delays for flights especially in harsh weather conditions in winter. The *BLC* was chosen because of the consensus that it is operating near capacity at some times.

The capacity will be modeled in the shortest time frame possible, from quarter of an hour to minutes. The short time frame for the capacity analysis is required because a lot of times there is a lot of unused capacity, but in a short time frame higher capacity is needed to ensure a proper operational outcome.

The effect of changing of the processes or infrastructure is not in the scope of this thesis. The processes are evaluated in the operating conditions prevailed at the study interval when the data was gathered and the processes assessed based on that information.

2. Literature review

Here concepts from the literature are introduced to structure the phenomenon called the airport. The nature of airport operations are being described from three perspectives: as a service supply chain, as interlinked processes, and in the framework of theory of constraints. These perspectives are needed to completely analyze the capacity of an airport and how the individual processes and services contribute to the capacity in a whole. Also a review is made on how the runway process capacity has been modeled in previous studies. The theories related to the thesis are introduced and explained. The section covers the relevant parts of service operations management and operations management in this thesis.

2.1 Services and processes

The airport is making possible flight and freight services to the passengers and other end-customers. Unlike in industry, no physical goods are produced at the airport. Airport operations are characterized by a large number of services being produced at the same time. There are many organizations creating these services and many customer-supplier relationships which create a service supply chain, an illustration of which at Helsinki airport was shown in section 1.1 (Research background – The operational efficiency of an airport) Figure 2. The services constitute larger processes that take place at the airport. Discussing the operations at the airport with these abstractions make it easier to understand the structure and nature of the operations that take place at an airport.

2.1.1 Service

What is a service? From the customer perspective, service is the combination of the customers' experience and their perception of the outcome of the service. There might be considerable overlap between the outcome and experience. (Johnston and Clark, 2008; p.8)

$$Service = experience + outcome$$

The customer experience is the customer's direct experience of the service process and concerns the way the customer is dealt with by the service provider.

The experience results in a set of outcomes: benefits, emotions, judgments (including perceived value) and intentions (Johnston and Clark, 2008, p.8). In an airport service supply chain network the experience is relevant to the end-customer receiving the flight service. The experience isn't that high a priority in many business-to-business service delivery relationships at the airport.

The service outcome describes the result for the customer of service delivery. In the case of an airport service supply chain network, the services provided for business customers like airlines the outcomes consist mainly of benefits as the value adding outcomes of the operations done by the service provider as they participate in the airport service supply chain network.

2.1.2 Relationships and roles of a service

Sometimes the end-customer can be seen as a stakeholder to another service taking place at an airport. The stakeholders of a service can be divided into payers, beneficiaries and participants (Johnston and Clark, 2008, p. 74-77). The end customer is in some cases a participant to the service and in some cases a beneficiary. The different roles mingle up in many services taking place at an airport.

For example the end-customer can be seen as a participant of the security check service at an airport. The end-customer is forced to be the participant of the service because of the laws and regulations constraining the aviation business. The actual customer, payer and beneficiary is the airport operating company and the service provider is the security company.

In customs the passenger is the customer fulfilling his or her duty as entering the country, but the payer, beneficiary and service provider is the state, in this case Finland of which customs is a government agency.

2.1.1 Service supply chain network

The end-customers at the airport are the individual consumers receiving the flight service to somewhere or the businesses buying flights or freight services. When the businesses buy flights for their employees, the employees are being the consumers of the service. As Lambert and Cooper (2000 p.66) suggested the supply chain of a company might not be a "chain" at all but a network of more complex relationships between suppliers and customer. So, the relationships of the organizations operating at an airport can be considered as a service supply chain network or more shortly service supply network, where different organizations provide services to each other and to the end customer.

The end customer is serviced by many service suppliers at the airport before receiving the final service themselves. As was illustrated in Figure 2 the relationships with and between the supplier organizations at Helsinki airport are manifold. This must be taken into account when trying to assess the capacities in different parts of the service supply chain network.

There might be service providers with different capacities in different processes. As many times service providers have contracts with only one airline, there might be capacity for one airline but not the other. It makes the assessing of capacity harder for the whole process. There could also be possibilities of global optimization of the capacities through contractual and operational changes between different customer supplier relationships. It might be possible to make better use of idle capacities of the suppliers with elevated fees in the benefit of the airlines whose service providers are running their operations at capacity.

Overall the complex relationships between the organizations operating at the airports service supply chain network make it hard to assess the capacity in many situations.

2.1.2 Services constituting processes

The services at an airport clearly constitute processes where there are inputs to a process and the processing results in outputs of the process. In this study a process model of the Helsinki airport operation is created to illustrate the complete model of the airport. Process model is good way to describe the whole process at the airport and it can be used in many ways to analyze and develop processes and to communicate how processes are related to each other.

Services can be seen as parts of processes. A **process** is a set of logically related tasks or activities performed to achieve a defined business outcome (Bozarth and Handfield, 2005, p. 46). The processes at an airport comprise of many activities, which constitute services to some customer. **Primary processes** address the main value-added activities of an organization. These processes are considered “value-added” because some customer is willing to pay for the resulting outcomes. **Support processes,** on the other hand, perform necessary, albeit not value-added, activities. (Bozarth and Handfield, 2005, p.46). In this study the main concern is on primary processes that directly have an effect on the succeeding in getting the passengers, baggage, freight and aircraft to the right place at right time and are thus value adding processes to the end-customer. The supporting processes are the ones that are making the value adding processes possible. The supporting processes are left out of the study.

Mapping business processes is the process of developing graphic representations of the organizational relationships and/or activities that make up a business process. It serves several purposes (Bozarth and Handfield, 2005):

- It creates a common understanding of the content of the process: its activities, its results, and who performs the various steps.
- It defines the boundaries of the process.
- It provides a baseline against which to measure the impact of improvement efforts.

In the later section 4.1 (A process model of the airport) process diagram of the Helsinki airport, is used to describe the whole process of airport as a one big process, with high level process view of the main processes.

Different types of graphical representations of processes have been devised. Some of as presented by Krajewski (2007 p.155) are:

- Flowcharts (i.e. process maps)
- Service blueprints
- Process charts

A Flowchart traces the flow of information, customer, equipment, or materials through the various steps of a process. Flowcharts are also known as flow diagrams, process maps, relationship maps, or blueprints. Flowcharts have no precise format and typically are drawn with boxes (with a brief description of the step inside), and with lines and arrows to show sequencing. (Krajewski, 2007, p.157)

The definitions of different presentations of processes vary from author to author and there isn't any standard way of depicting many of them. Some standardized process presentations are devised like in Unified Modeling Language (UML), but none of them seemed to serve the purposes of this study so later in the section 4.1 (A process model of the airport) a hybrid model of a process flowchart and a material flow diagram is used to describe the primary processes at Helsinki Airport.

2.2 Theory of constraints (*TOC*)

> *"What is important to notice is that the prevailing notion that 'more is better' is correct only for the constraints, but is not correct for the vast majority of the system elements – the non-constraints. For the non-constraints, 'more is better' is correct only up to a threshold, but above this threshold, more is worse. This threshold is dictated by the interdependencies with the constraints and therefore cannot be determined by examining the non-constraints in isolation. For the non-constraints, local optimum is not equal to the global optima; more on the non-constraints does not necessarily translate to better performance of the system." (Goldratt 2010, p.4)*

The theory of constraints (*TOC*) is a systematic management approach that focuses on actively managing those constraints that impede a firm's progress toward its goal of maximizing total value-added funds (Krajewski, 2007, p.254). The theory was introduced by Eliyahu M. Goldratt and Jeffrey Cox in their book "The Goal" (Goldratt and Cox, 1993). In this chapter some of the key properties, concepts and methods associated with the theory when considering capacity of a service system are introduced.

The key idea behind the theory of constraints is, as described by the quote from Goldratt's and Cox's book, that some processes are the constraining ones for a system as a whole, and those processes set the capacity for the system. If there were no constraints the capacity of a system would in effect be infinite. Goldratt and Cox were describing the situation mainly in manufacturing context, but in can be as well be used in service context. In an airport this means that given a situation, some processes are constraining the system as a whole, making it impossible to operate faster than the constraining processes determine. On the other hand in the specific situation adding resources to the non-constraining processes don't improve the overall system performance.

2.2.1 Constraints

A constraint is any factor that limits the performance of a system and restricts its output. Constraints can occur up or down the supply chain, with either the firm's suppliers or customer, or within one of the firm's processes like service/product development or order fulfillment (Krajewski, 2007, p.254). In an airports service supply chain network the constraint can be observed in multiple of places given the current state of the airport. The current state of the airport is a reflection of its past events and disturbances in operations can accumulate a situation where normally non-constraining resources become constraining resources.

The principal tenet of *TOC* is ''constraints determine the performance of a system''. Since there are few constraints in any system, management of these few key points allows for effective control of the entire system (Watson et al., 2007, p.391). *TOC* suggests that the constraint should be exploited to get the maximal capacity from the system. Exploitation of the constraint seeks to achieve the highest rate of throughput possible within the confines of the system's current resources. The output of the system is limited by the rate of throughput at the constraint (Watson et al., 2007, p.391). This is why according to the *TOC* the whole system should be run in the pace of the constraint and the constraint be working at 100% capacity at all times to achieve the greatest possible throughput of the system.

There are generally three types of constraints identified in the literature: physical (resource capacity less than demand), market (demand less than resource capacity) and policy of managerial (formal or informal rules that limit productive capacity of the system) (Watson et al., 2007). This study is focused on the physical constraints of the system where the capacity of process is constraining the airport operations because it resources don't allow it to perform any faster.

2.2.2 Capacity

The definition of capacity is not always clear at all and it varies from process to process. Many times the capacity is a measure of something in a time interval like number of baggage in an hour or aircraft per hour, but it can also be a constant value like in integrating processes like a warehouse, which has the ability to hold a certain amount of physical goods, but without any time constraint.

Capacity is the maximum output of a process or a system. The capacity can in most cases be measured somehow. No single measure is best for all situations. In general, capacity can be measured in two ways: in terms of output measures or input measures (Krajewski, 2007, p.255). The output measures of capacity are best utilized when applied to individual processes within the firm, or when the firm provides a relatively small number of standardized services and products (Krajewski, 2007, p.256). In an airport context most of the capacities can be measured in measures that describe number of outputs per hour, because most of the processes are processing some tangible objects. On the other hand some processes are storing objects and time related capacity measure isn't meaningful, like in baggage hotel, which stores the baggage brought long before departure, before it is sent to the aircraft.

The input measures are generally used for low-volume, flexible processes (Krajewski, 2007, p.256). In airport context the input to the process is many times the same as output like the security check or the deicing process of an aircraft, so it is the same thing to use the input measure or the output measure as the capacity measure.

Theoretical capacity or the design capacity is the absolute maximum capacity that a system or a process can handle. It is the maximum output capability, allowing for no adjustments for preventive maintenance, unplanned downtime, or the like (Bozarth and Handfield, 2005, p.210). It is many times the same as design capacity because the capacity is constraint by other factors than what have been taken into account at design time of the system. On the other hand practical capacity or rated capacity, is the long-term, expected capability of a resource or a system (Bozarth and Handfield, 2005, p.210).

2.2.3 Utilization

Utilization is the degree to which equipment, space, or the workforce is currently being used, and is measured as the ratio of average output rate to maximum capacity:

$$Utilization = \frac{Average\ output\ rate}{Maximum\ capacity} *\ 100\%$$

Utilization is a very useful measure in an airport context in relation to the *TOC*. When the utilization of a process' capacity is near 100% it may become a constraining for the whole system, the airport. In this type of situations it might be beneficial increase the maximum capacity if it is possible. The decision comes then of course to the cost-benefit analysis of increasing the capacity.

2.2.4 Bottleneck

Bottleneck is a special type of a constraint that relates to capacity shortage of a process, and hence is also referred to under certain conditions as a *capacity constraint resource (CCR)*. It is specifically defined as any resource whose available capacity limits the organization's ability to meet product volume, product mix or demand fluctuation required by the marketplace. A business system or a process would have at least one constraint or a bottleneck; otherwise its output would be unlimited. (Krajewski, 2007, p.254)

Bottlenecks can both be internal or external to the firm, and typically represent a process or a step with the lowest capacity and longest **throughput time**, which is the total time taken from the start to the finish of a process. Where a bottleneck lies in a given service or manufacturing process can be identified in several different ways. The bottleneck could be occurring at the workstation with the highest total time per unit processed, or the workstation with the highest average utilization and total workload, or the workstation where even a single minute reduction in its processing time would reduce the average throughput time for the entire process. (Krajewski, 2007)

2.2.5 *TOC* Five step focusing process

Theory of constraints (*TOC*) holds that improvements in the systems performance can only be achieved by managing the constraints. *TOC* includes a five step focusing process, which draws the attention to the key issues in order to improve the systems output. The focusing process is managerial in a sense that it merely guides the efforts of the management to the right place in the pursuit of a system performance improvement. The process steps are (Siha, 1999, p.256):

1) Identify the system constraint(s). A system cannot be maintained at maximum performance unless we know what constrains the system so we can design control mechanisms appropriate to the constraints.
2) Exploit the system constraint(s). We must make the best possible use of the constraints. For example, physical constraints within the system must be scheduled to produce the most profitable products.
3) Subordinate the non-constraint(s). Non-constraints, by definition; do not limit maximum performance of the system. Decisions affecting constraints must take priority over those affecting non-constraints.
4) Elevate the constraint(s). After completing the above steps, further improvements in performance of the system require changing a constraint. Increasing the capacity of a machine that constrains profit is an example of this step.
5) Return to step 1. After a constraint is changed, new system constraints may surface. Return to step 1 to identify new constraints.

In this study the aim is to simply try to identify in what kind of situations certain processes become system constraints or bottlenecks by modeling the system. In that sense the focus is on the first step of the focusing process.

2.3 Capacity of a runway

Airports are studied in the field of transportation research. A lot of research has been done in modeling the capacities of the runways. The relevant theory and findings on runway capacities are presented here. Later the capacity of the Helsinki airport is studied empirically.

Runway capacity is defined in aviation research as “the maximum sustainable throughput of aircraft operations; both arrivals and departures that could be performed during a specified time interval (e.g. 15 minutes, or an hour) at a given airport of a specific runway configuration, under given weather conditions and at acceptable level of aircraft delay”. (p.237 Ashford et al., 2011). In this thesis the capacity of the runways studied as the function of the mentioned conditions.

The runway capacity can be estimated by measuring the interoperation times from operational recorded and observation at busy airports. Analytical methods can be used to estimate capacity based on interoperation time, which in turn is influenced by the stochastic variability in aircraft speeds, variation in runway occupancy (caused by variations in aircraft performance characteristics), and other operation base factors (p.239 Ashford et al., 2011). In section 4.2 Runway process, the capacities are estimated based on the operational database records.

There are a lot of factors affecting the runway capacity of an Airport. Factors that influence runway capacity according to Ashford et al. (2011) include:

• Meteorological conditions in terms of visibility, cloud ceiling, and wind

• Airfield layout, runway configuration, and operational strategy of using the runway at different wind directions

• Aircraft arrival and departure ratios

• Aircraft fleet mix as related to approach and departure sequencing, and runway occupancy time per aircraft type

• Runway occupancy times as related to aircraft performance characteristics and runway exit location

• *ATC*-related matters in relation to runway arrival fix loading, sector loading, ATM procedures during congestion times, and controllers’ work load

In this study some of these factors’ effects on the case airports runway capacity are studied.

In the research on airport and airfield capacities it has been established, that the arrival and departure capacities are connected with each other through a convex, nonlinear functional relationship (Gilbo, 1993, p.145). This idea is presented for example by Gilbo (1993) and Ashford et al. (2011). This relationship can be presented by equation,

$$C_d = \phi(C_a),$$

where C_d is departure capacity and C_a is arrival capacity. ϕ in turn is a nonlinear, concave decreasing function. The function ϕ depicted in Figure 3. The figure represents a capacity curve of a under certain conditions. To make the relationship specific for an airport requires a complex approach that includes a combination of mathematical modeling using empirical data, and validation of the results using the expertise practicing traffic managers and controllers (Gilbo, 1993).

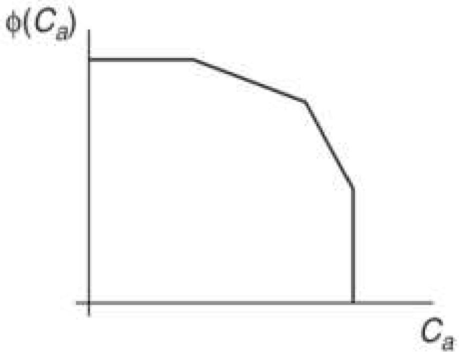
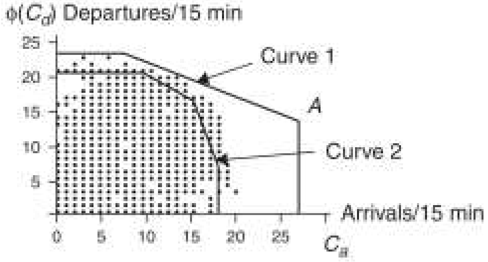


Figure 3: Airport arrival/departure capacity curve (Ashford et al., 2011; Gilbo, 1993)

Gilbo (1993) presented an empirical method for estimating and presenting this airfield capacity curve. The method is based on the assumption that during the considered time the peak arrival and departure counts reflect the airport performance at near capacity level. Capacity curves could be estimated for specific set of conditions from the observed data.

Gilbo's method consists of fitting a piecewise linear curve around the observed arrival/departure per 15 minutes combinations. This same procedure is done for all the relevant set of conditions to be studied. The application of this method can be seen in the Figure 4 and Figure 5: Capacity curve estimation method from frequencies. (Ashford et al., 2011; Gilbo, 1993).



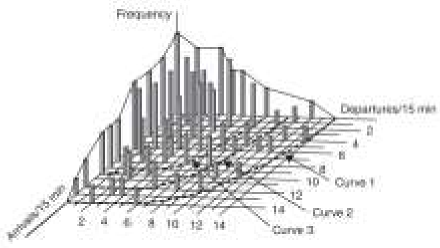


Figure 4 and Figure 5: Capacity curve estimation method from frequencies. (Ashford et al., 2011; Gilbo, 1993)

There can be outliers in the empirical data which can be problematic. There can be kinds of outliers: errors in the data collection or they can reflect real but rare events an airport operates beyond it normal operational limits. Gilbo tackled this by using a frequency based point rejection criteria and calculating a confidence levels on capacity estimates based on this.

Ramanujam and Balakrishnan 2009 devised Gilbo's method using piecewise linear quantile regression methods for the airfield capacity curve estimation. The quantile regression was modified to include constraints to the estimation of the piecewise linear quantile regression curve. The constraints included the concavity and non-positivity of the piecewise linear slope estimates. This method was applied to data from the New York area multi-airport system comprising of Newark, John F. Kennedy and LaGuardia airports. Ramanujam and Balakrishnan demonstrated in their research that their quantile regression based statistical technique enabled the identification of the impact of the key factors influencing the capacity envelopes. Quantile regression based technique was found to be a good method for systematic outlier elimination from the data.

Later in the findings chapter, section 4.2 Runway process, a piecewise linear quantile regression is used, in a similar way as in the presented studies, in estimating the runway capacity in different scenarios at Helsinki Airport. A statistical language R procedure was used in the estimation.

3. Methodology

3.1 Statistical methods used

In this section the used statistical methods are introduced as well as short introductions to the theory behind these methods in the context of this study.

3.1.1 Linear Quantile Regression

A distribution of a variable can be divided into fractiles. Median is specifically the 50% point of the variable. Half of the values of the variable are above, and half below this value. Similarly quartiles include the information of distribution of a variable into four equally sized proportions of a variables distribution. Even smaller division of a distribution is the division to percentiles which divide the distribution into 100 proportions. A variable can be further divided in to quantiles which are the generalization of percentiles, quartiles and all other fractions of observations of a variable.

Quantile regression is the act of estimating conditional quantile functions in which quantiles of the conditional distribution of the response variable are expressed as functions of observed covariates (Koenker and Hallock, 2001, p.1) This is in contrast to, for example, standard linear regression where the purpose is to estimate the expected value of the distribution. In quantile regression the regression model is explaining the variation in a distributions quantile τ.

Quantile regression makes it possible to model how any quantile of the response variable behaves when the independent variables change. In some situations a certain quantile of the response variable might be interesting, e.g. if it behaves differently in some percentile or when trying to define the upper bound or the lower bound of a response variables distribution. In the case of capacity it’s very interesting to know were the upper bound of the load distribution is. By studying the load distributions near maximum values it’s possible to say what the maximum capacity in practice is (if the studied capacity in question is at least some times operated near maximum values).

The linear quantile regression model the conditional quantile function $Q_y(\tau|\boldsymbol{x})$ (conditional to the covariates x values) is estimated. The conditional quintile function for quantile τ is the dependent variable Y. The regression model for the quantile τ is of the form,

$$Q_y(\tau|\boldsymbol{x}) = Y = \boldsymbol{x}'\boldsymbol{\beta}(\tau \,|\, \boldsymbol{x}) + \varepsilon$$

where ε is the error term or residual, x is the vector of the covariates values, and $\boldsymbol{\beta}(\tau \,|\, \boldsymbol{x})$ are the τth quantile coefficients. The model explains how the τth quantile of the distribution behaves as a function of the covariates.

In the same way as the mean of the response distribution in standard linear regression can be estimated by minimizing the sum of squared residuals of the model, it is possible to estimate the median of the response variable by minimizing the sum of absolute values of residuals of a regression model (Koenker and Hallock, 2001). The formula of absolute value of residuals can be weighted with different weights corresponding to the desired quantile to be modeled τ represented by a number between 0 and 1. When the weights τ and $\tau - 1$ are chosen so that they sum up to 1 and $0 < \tau < 1$. The τ:th quantile function can be estimated by minimizing the sum of weighted absolute residuals.

$$\underset{\boldsymbol{\beta}}{min} \sum \rho_\tau(y_i - \boldsymbol{x}_i^{\mathrm{T}}\boldsymbol{\beta})$$

Inside the parentheses are the absolute residuals $y_i - \boldsymbol{x}_i^{\mathrm{T}}\boldsymbol{\beta}$. The loss functions $\rho_\tau(u) = \left| u\left(\tau - I_{y<0}\right) \right|$ weighs the positive and negative residuals by τ and $\tau - 1$ respectively. An illustrative picture of the asymmetrically weighing function $\rho_\tau(u)$ is shown in Figure 6.

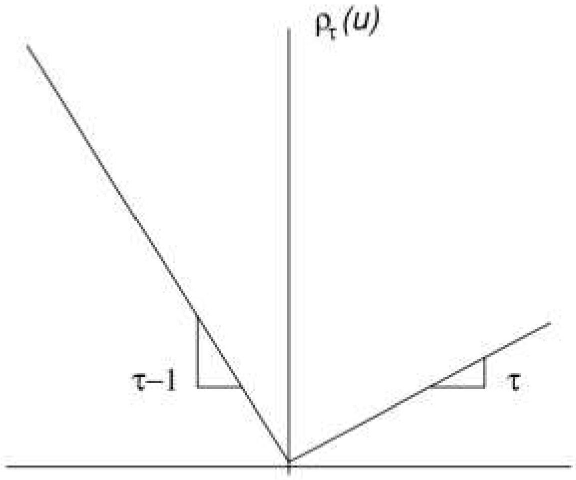


Figure 6: Quantile regression ρ function. (p.6, Koenker, 2005)

The coefficient estimates $\hat{\beta}_\tau$ for quantile τ can be computed by solving

$$\hat{\beta}_\tau = \underset{\boldsymbol{\beta} \in \Re^k}{\operatorname{argmin}} \sum_{i=1}^{n} \left(\rho_\tau(y_i - \boldsymbol{x_i}'\boldsymbol{\beta})\right)$$

$$= \underset{\boldsymbol{\beta} \in \Re^k}{\operatorname{argmin}} \sum_{\{i \mid y_i - \boldsymbol{x}_i'\boldsymbol{\beta} > 0\}} \tau \left| y_i - \boldsymbol{x_i}'\boldsymbol{\beta} \right| + \sum_{\{i \mid y_i - \boldsymbol{x}_i'\boldsymbol{\beta} < 0\}} (\tau - 1) \left| y_i - \boldsymbol{x_i}'\boldsymbol{\beta} \right|$$

Where the coefficients β are chosen that minimize the sum of weighted absolute errors of the regression model. Finding the solution to this problem requires linear programming techniques.

The linear quantile regression estimates can be done for different subsets of the independent variables when the different subsets are fitted with their own quantile regression lines creating a piecewise linear hyper surface or a curve in two dimensional case. This type of piecewise quantile regression can be used in situation where there is non-linearity present.

Later in section Runway process, a piecewise linear quantile regression model is used to model the runway capacity in different scenarios and conditions.

3.1.2 Generalized linear models: Gamma regression

Generally the purpose of statistical model fitting is trying to understand in broad terms the functionality of the underlying phenomenon. The model is a simplification of reality, but it is able to describe and give information of the functional relationships present in the object of the research.

In this section the statistical regression technique called the gamma regression is introduced. Gamma regression is used for example in modeling inter-arrival time problems, survival time problems, and such (Myers, 2010). Generally gamma regression is used mostly on settings were the duration of something (i.e. time) is modeled. Later in the section 0 The output of the deicing process is the deicing treatment done (the provided service) to the plane. Its capacity is measured as treatments done per unit of time. It is clear that the time taken to process on plane is inversely proportional to the number of planes that can be treated in an hour in a single processing line. So the expected Deicing capacity C_{DI}, can be expressed as,

$$C_{DI} = \frac{1}{E(t)}[treaments/\mathrm{hour}] \tag{1}$$

where $E(t)$ is the expected time taken to process a single plane in hours.

There are different conditions or factors that affect the treatment time of single plane. Let A, B, C be the set of all the different combinations of conditions that have known effect on the treatment time.

$$A, B, C, \dots are\ sets\ of\ categorized\ conditions\ for\ a\ deicing\ treatment$$

The set A can be for example the types of aircraft or categorized average wind speeds. K is a set all the possible combinations of the discreetly categorized condition sets,

$$K = \{(k_A, k_B, k_C \dots) \mid k_A \in A\ k_B \in B, k_C \in C \dots\ \}$$

I is an index set for all the elements of K

$$\left\{ i \in I \mid K = \bigcup_{i \in I} K_i \right\}$$

If the conditions can be forecasted for the next day and weights w_i can be assigned to each condition combination $\in I$. Then w_i are the weights of proportionally how many aircraft are treated in conditions $i \in I$ (i.e. w_i is how many planes are to be treated in conditions i when divided by the total number of planes to be treated).The theoretical total capacity of deicing treatments for a single processing line can be calculated with formula (if it can be assumed that the next treatment can start straight away after the one before):

$$C_{DI} = \frac{1}{\sum_i w_i E(t_i)}, \qquad \sum w_i = 1, \tag{2}$$

The different processing lines can be marked with $j \in \{1..N\}$, where N is the number of processing lines. When there are more than one processing line and the weights w_{ji} for different conditions K_i are different for the processing lines j, the expected total capacity C_{DI} can be expressed as:

$$C_{DI} = \sum_j \frac{1}{\sum_i w_{ji} E(t_i)}, \quad \sum w_i = 1 \tag{3}$$

In order to be able to calculate the total capacity the expected treatment times t_i need to be estimated for the set of condition combination $K_i,\ i \in I$. In the next section a model for estimating the expected value of deicing times is introduced.

Deicing processing time model, a gamma regression model is fitted in order to describe the functional relationships between different conditions and the deicing treatment processing time, hence producing a link between the capacity and the conditions.

Generalized linear models are a class of regression models that are related to the exponential family of distributions. Gamma regression model is generally applied in a situation where the modeled response variable is positive, continuous and its variance is directly proportional to the square of the expected value (Myers, 2010). This means that the when the expected value increases so does the variance of the dependent variable. It is in contrast to the linear model where, the variance is expected to be constant (homoscedasticity).

A gamma regression model of the form,

$$Y_i \sim Gamma(\mu_i, \nu)\,, \qquad g(\mu_i) = \eta_i, \qquad \eta_i = \boldsymbol{x_i}'\boldsymbol{\beta}, \qquad i = 1,2, \dots n$$

Where Y_i is the dependent variable, which is gamma distributed, and there are n in total of observations. The link function $g(\cdot)$ is relating the expected value of the observations of the dependent variable $E(y_i) = \mu_i$ to the linear predictor η. The observation vector $\boldsymbol{x_i}$ is the independent variables (predictor variables) and $\boldsymbol{\beta}$ is the regression coefficients vector of the predictor variables. The linear predictor $\eta_i = \boldsymbol{x_i}'\boldsymbol{\beta}$ with the link function $g(\cdot)$ determine the expected value of an observation i, as the link function $g(\cdot)$ is smooth an invertible:

$$E(y_i) = \mu_i = g(\eta_i)^{-1} = g(\boldsymbol{x_i}'\boldsymbol{\beta})^{-1}$$

When the link function is identity i.e. $g(\mu_i) = 1 * \mu_i = \mu_i$, the expected value is the same as in the standard linear regression

$$\mu_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \cdots = \boldsymbol{x_i}'\boldsymbol{\beta}$$

The variance of a gamma regression dependent variable is,

$$V(Y) = \phi\mu^2$$

where ϕ is the dispersion parameter of the estimated regression model (p. 383, Fox, 2008). The dispersion parameter is the constant in the model relating linearly the variance of predicted variable Y to the expected value μ of the dependent variable. The variance is depicted as a function of the predicted variables value, because the variance is not constant. The dispersion parameter is constant in the whole model, so the only thing affecting the variance of the dependent variable at a given point is the expected value $E(Y)$ of the explained ble Y.

The main difference between the gamma regression with identity link and a standard linear regression is that the error term ε_i of each observation is gamma distributed and not normally distributed.

$$y_i = \boldsymbol{x}_{\boldsymbol{i}}'\boldsymbol{\beta} + \varepsilon_i$$

Also the variance of the gamma distributed error term ε_i increases as the expected value increases in the way described in the previous paragraph. These properties make the gamma regression much more suitable model for many duration estimating models than the standard linear regression.

3.2 Data collection

The data collection methods used ranged from informal methods like discussions with informants to gathering data from the case airport operator company's (Finavia Oyj) databases that are shared with the airlines operating at Helsinki-Vantaa and operational databases of Finnair Oyj.

For the complete airport process model Finnair Oyj company's process flow charts of their processes were used as information sources. These charts illustrate the processes in Finnair's point of view at the airport. The charts were used with discussions with informants at the airport to create a process model diagram of the Helsinki Airport.

The source for the airfield arrival and departure data is the Scope database used by Finnair to which the airport operator Finavia provides the data. The data had to be heavily modified to get it to the right form.

The source for the deicing times was the CAPCO database and software which records all the deicing activities at Helsinki-Vantaa airport for all the stakeholders and deicing service suppliers.

The baggage process information was gathered from the *BLC*'s information systems as well as discussions with its employees.

Weather observation data from Finnish Meteorological Institute's open weather observation catalog (Finnish Meteorological Institute, 2013) is used in the runway capacity analysis as well as the deicing process analysis. The weather data used in the study were taken from the Helsinki-Vantaa Airport weather observation point. Because some snow depth measurements were missing for approximately 10 days, they were substituted with weather observations from Kumpula, Helsinki observation point. The substituted days were actually winter storm days during which there was a lot of snowfall.

4. Findings - Modeling the airport processes

First in section 4.1 A process model of the airport, a schematic model of Helsinki Airport and its processes as system are created to give an abstract view of the processes and operations that are interlinked and have to be taken into account when considering the capacity of the airport. Then in section 4.2 Runway process, the capacity of the runways at Helsinki Airport are analyzed as well as how some factors have affected the capacity. In section 4.3 Deicing process, the deicing process capacity is analyzed and finally in section 4.4 Baggage process capacity – Baggage Logistics Centre the capacity of baggage processing is analyzed.

The used models and techniques in analyzing the capacities, utilization and determining the effects of some factors on the processes are described in detail. Also the validity and the reliability of the analyses are considered as well as possible error sources are mentioned.

4.1 A process model of the airports primary processes

An airport can be viewed as a big process where the inputs (arriving aircraft with passengers and baggage, local departing passengers and their baggage) are turned into outputs (arriving local passengers, baggage and departing aircraft). Helsinki airport is depicted in this way in a diagram in Figure 7. The diagram depicts Helsinki Airport as a system with its main flows of objects: passengers, baggage (but excluding freight), and aircraft. The flow of the freight was left out because of its different nature and additional information gathering requirements. The diagram was composed with information gathered from informants and an introductory tour around the Helsinki Airport with Duty Manager, HUB control center, 2013. In addition, Finnair Oyj company's process flow charts of the processes in Finnair's point of view at the airport were reviewed.

The diagram in Figure 7 is not a standardized diagram. It has some features from a process flow chart as well as from a material flow chart. Whereas a process flow chart depicts the processes in a workflow and time perspective, Figure 7 represents processes in physical flow of objects and time perspective. In a process chart all the processes are depicted as logical diagram of what process needs to come after the previous. The object of the process can change from managing a passenger (a customer) to managing a physical location at the airport. So the difference is the perspective from the work needed to be done. In Figure 7 each arrow represents either a physical movement or a state change of the flow, from the process or state the back end of the arrow is connected to, to the process or state the front end of the arrow connects to. Each process in Figure 7 also represents a physical location where the process is conducted. For capacity consideration it is significant to depict the physical locations as well. For instance the numbers of gates actually contribute directly to the maximum number of aircraft at the gates and hence the capacity of the gate processes. The same goes for the security check stations and check-in counter and baggage drop processing locations. The logic of going from one process or state to another is also left out from the diagram because it would have made it visually complicated.

The diagram represents a fairly high level illustration. Each of the processes could be divided into smaller processes or activities. Not all the aircraft, passengers and baggage are processed the same way in the main processes. All the main flows are depicted while some have been left out because of insignificance and some minor irregularities might prevail. Nonetheless the main aspects of the processes and main flows of objects at Helsinki Airport affecting its capacity are evident.

From capacity point of view, the arrows depicting the flow of objects from one location or state to another are also important. The flows themselves might have some capacity or maximum throughput associated with them as well as throughput time. For example as the passengers move around the terminal the physical sizes of the passageways constrain the flow of passengers from one point to another. Also as the aircraft are taxiing at the terminal airside, there are limits to how many of them can be moving at the same time and it is depending on multiple factors. The capacity of the single arrows depicting the movement of aircraft from one place to another are themselves complex tasks to evaluate.

The “storage” states (waiting/stock/queue) of the object flows depicted in the diagram include many events and possibly activities in them. For example the Helsinki-Vantaa terminal is depicted as two different “storage states” of passengers: Non Schengen area and Terminal 1 & 2 gate area. In reality the passengers are moving around the terminal, doing many things and possibly involved in many service events. Another example is the maintenance state of the aircraft, which from the airport system point of view can be regarded as storage, but actually includes the maintenance processing as well. The capacity considerations should also be extended to these states.

The movements of arriving and departing aircraft are divided into two in the diagram of Figure 7. This is has been done because the flows of objects are different for the passengers, to reflect the difference of their physical location in the terminal buildings and the fact that a non-schengen flight cannot arrive at or depart from the schengen area or the other way around. This way the schengen non-schengen division also affects the capacity available for a flight, because an aircraft must always use the appropriate resources of the terminals due to the regulations.

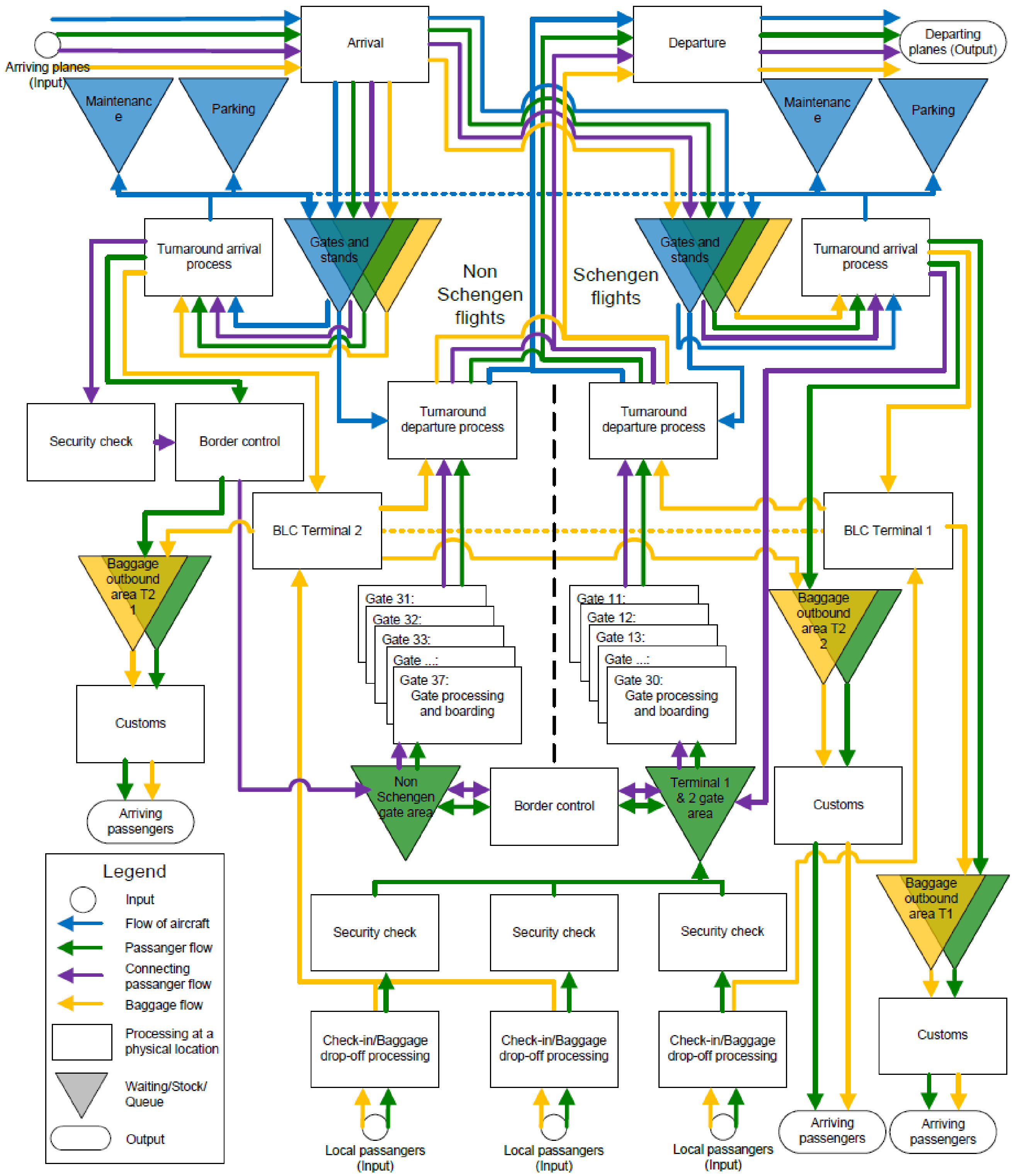


Figure 7: The Helsinki Airport depicted as a system of processes with input, outputs and the main flows of objects: passengers, baggage and aircraft

Any distraction in the flows of objects (aircraft, passengers, baggage) can restrain the “objects” from getting together at the right time and at the right place in effect creating a bottleneck for the process, and possibly create a delay for a flight. Unlike in a factory the flows in an airport don't always go through predefined routes and phases, but it constitutes an open system where everything affects everything. The relationships are so manifold that in reality the process capacities are complex function of inputs, outputs, resources and time. In contrary a factory can in most cases give very precise figures of how many products can be produced in a day or how many components can be processed at a given process in a given time. The situation is much more controlled. At an airport the flows are in different magnitude of randomness.

A striking observation is the complexity of the airport as a system of processes. As the purpose of this thesis is to analyze the capacities and forming of bottlenecks at Helsinki Airport and it is quite clear how complex the task is when considering the processes even in a high level diagram. A bottleneck can occur in the airport system in multiple of ways and at any point of the diagram. For instance when an aircraft arrives there needs to be enough capacity in the runway for the arrival. Otherwise the aircraft needs to keep flying until capacity is available, which will take time and possibly delay the later processes. When the aircraft has arrived there needs to be enough capacity in the taxiways for the aircraft to move to the gate that has been assigned to it. When the plane arrives at the gate the turnaround around arrival process starts. The turnaround arrival needs to have enough resources available, i.e. capacity, to process the plane in order to not delay the next stages the aircraft goes through. The divergence from the planned schedule of the activities related to the turnaround arrival process or any other process don't necessarily cause the overall schedules to fail. If the same aircraft is directly scheduled to another flight from the same gate, there needs to be enough capacity at the gate to process the boarding passengers and at the same time the turnaround departure process needs to have the capacity to timely manage the processing of the departing aircraft. When the aircraft has finished its turnaround departure process it starts taxiing' to the runway and then again the taxi ways and runways need to have the capacity in place in order not to cause delay.

Due to the above-described complexity, this study represents few processes that can become bottlenecks in certain situations, which were pre-identified with the study initiator company. The processes that are analyzed in the following sections can be seen in Figure 7. The arrival and departure processes i.e. the runway capacities are analyzed in section 4.2 (Runway process). The capacity of deicing process, which is part of the turnaround departure process, is analyzed in the section 4.3 (Deicing process), the baggage logistics centre is analyzed in section 4.4 (Baggage process capacity – Baggage Logistics Centre).

4.2 Runway process capacity

Runway capacity can be seen as a process producing two outputs: arrivals and departures. These two outputs restrict each other so there is functional relationship between the arrivals and departures. As described in section 2.3 Capacity of a runway, this relationship can be seen as departures as a function of arrivals. The arrivals generally take more often precedence over the departures as it is easier for the aircraft at the airport to wait on the ground than for the arriving aircraft in the air.

There are three runways at Helsinki-Vantaa airport (the international code for Helsinki airport is *EFHK* by international Civil Aviation Organization's *ICAO* standard). The three runways and their relative positioning at *EFHK* can be seen in Figure 8. Based on the current conditions like weather, airfield condition, arrival and departure schedule, regulations etc. the Air Traffic Control chooses the arrival and departure runways and direction for the arriving and departing flights. The runway and direction combination for arrivals and departures used at a specific time is called a runway configuration.

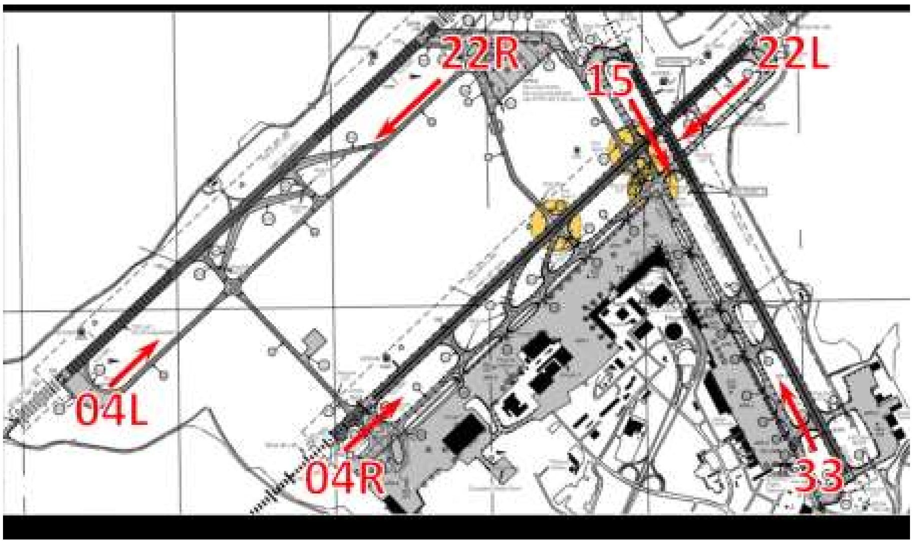


Figure 8: *EFHK* Helsinki-Vantaa airport runways. (Finavia Oyj, 2013b)

The *ATC* controls the usage of the runways. Most of the time the runways are used in the same configuration i.e. same arrival and departure runways for long periods of time until the weather changes. Change in weather is the strongest reason to change the runway configuration in use. Also the most preferred runway configurations are to use runways 22L and 22R or 04R and 04L simultaneously for arrivals and departures. The reason for this is that when using these configurations the arriving and departing aircrafts are moving to the same direction when airborne. The airspace is easier to manage that way and more actions are allowed by the operating regulations set by Trafi which is the civil aviation regulatory authority in Finland.

4.2.1 Modeling the capacity of runway operations

As described in section 2.3 the runway capacity can be estimated from empirical data and expressed as a piecewise linear curve between the arrivals and departures in a time interval. Also was introduced how quantile regression based technique can be used to estimate the capacity envelope of a runway configuration. A short theory on quantile regression relevant to this study was introduced in section 3.1.1.

The runway capacity at Helsinki-Vantaa airport is modeled here with piecewise linear quantile regression. A quantile regression package “quantreg” created by (Koenker et al., 2013, Koenker 2013) of the open-source statistical software R is used in estimating the capacity curves. The “rgl” package (Adler and Murdoch,

2013) of R is used in plotting. R software or its package “rgl” don’t have ready procedures to plot 3D histograms. So the scripts for 3D histograms had to be devised based on the hist3d() and binplot.3d() demo functions in the “rgl” package. The scripts created to plot the 3D histograms are in Appendix A: R – function: hist3d() for 3D-histogram plotting and Appendix B: R – function: binplot.3d() for 3D-histogram plotting.

4.2.1.1 Data

The data for the runway capacity modeling was gathered from the Finnair’s SCOPE database which uses flight data provided by the airport operating company Finavia. The database has information on all the flights taking place at the case airport Helsinki-Vantaa. The data represents all the flights arriving or departing from Helsinki Airport and details such as arrival or departure runway, the airline designator, aircraft type, timestamps of different events and so forth. The used fields where the touch-down time for the arriving aircrafts and take-off time (the time when wheels stop touching the runway) for the departing flights.

The data has 9680 15 minute intervals that had runway actions (arrivals or departures) between 1.11.2012 and 31.3.2013. The data isn’t necessary large enough to reliably estimate the effects on capacity of some factors that happen on occasions with low observation count, but it still gives reliable “lowest boundary” of maximum capacity. The real maximum capacity cannot be below the estimated capacity curves introduced in this section.

The operational runway data was combined with weather data from the Finnish Meteorological Institutes hourly weather data (Finnish Meteorological Institute, 2013). All the 15 minute observation points were combined with the weather data of their respective running hour, which was aggregated to an hourly level. This causes some inaccuracies in the estimations.

The arrival and departure runways were not available for other airline designators’ flights than Finnair and the information could not be retrieved on those flights. This problem was circumvented by calculating the mostly used runways for both arrivals and departures for Finnair’s flights for all hours during the study interval. It is assumed that the arrival and departure runways are the same for the rest of the flights arriving or departing by other airline designators during the same hour. This is a reasonable assumption because the runway configuration changes quite rarely. The error occurring in maximum capacity estimation due to this is minimal near the maximum values and 100% quantiles as described in section 2.3 Capacity of a runway.

4.2.1.2 Error sources in data

Here the identified problems and error sources in the data are presented and their effects' magnitude evaluated mostly on qualitative basis as quantitative measures are impossible to assess.

The deduction of runway combination for each flight operation was based on the assumption that the mostly used runway combination during each hour is the runway configuration used during that hour is the same for all flight operations (arrivals and departures) taking place at the hour in question. The estimated runway configuration in use was based on the flight operations of Finnair at that hour. The reason for this is that data on the used runways wasn't available for the other airlines than Finnair. Because the assumption made on the used runway configuration, around 6% of the flight operations of Finnair in all the observations (15 minute time intervals) are known to be in wrong runway combination and for sure some of the other airline's operations are counted in a wrong runway configuration as well since their flights operating runways are based on the Finnair's flights.

In the situations where it seems that multiple runway configurations have been in use by Finnair flights, there has most likely happened a change of the runway configuration in use, or there has been so little runway operations that it hasn't mattered which runways to use and in which direction. Nevertheless, this may have an impact on the capacity curve estimation making the capacity estimates little bit more optimistic. On the other hand, that is one of the reasons why 99.9%th quantile is used in maximum capacity estimation, which leaves out the clear outliers from the capacity curve estimation.

In each observation (15 min time interval) the other airlines' flights are presumed to have the same runways as Finnairs' flights, which is not always the case. The runway combination might have changed during a time slot, but the rest of the flights are other airlines' flights so they have been wrongly assumed to have operated in the previous runway configuration. This may have some impact on the estimation but its magnitude is hard to assess.

Sometimes the runways can be switched on the fly between arrivals and departures or the both arrivals and departures are operating from a single runway. If first one happens they might make the observations seem like a single runway arrival and departure configuration. That's why the single runway configuration observations are more unreliable than others.

The weather data is hourly data so it is assumed that all the 15 minute intervals have the same meteorological conditions as the hour they are included in. The reality on the other hand can be that the all the 15 minute time intervals during an hour might not be the same in terms of weather conditions. This has some minor effects on the results as there probably are very low amounts of these points where the weather conditions between consecutive 15 minute intervals are considerably different.

4.2.2 Planned load and utilization

The air traffic at Helsinki Airport is heavily crooked at the peak periods unlike in many greater airports such as Heathrow in London or Frankfurt in Germany. By this is meant that there are times when there are a lot of arrivals and almost no departures or the other way around. At big hub airports the arrivals and departures are more evenly distributed, because there's more traffic to all directions at all times of the day. At Helsinki Airport the crooked structure of the runway utilization is due to the location of Helsinki being very beneficial location when travelling between central Europe and Asia. Helsinki Airport is at nearly optimum location for connecting flight in terms of distance needed to be travelled.

In principal the European flights bring customers to Helsinki Airport that connect to Asian flights. The European flights and the Asian flights arrive at Helsinki in the afternoon, they make a turnaround, take new passengers and take-off. The Asian flights go back to Asia, but the other flights might go to other destinations.

The flight traffic at Helsinki Airport has two greater peaks or banks in a day: in the morning from around 6 a.m. to 10 a.m. local time and in the afternoon from around 2 p.m. to 18 p.m. local time. The planned flights, average arrivals and departures, per 15 minutes at Helsinki Airport are depicted in the Figure 9. The maximum and minimum values are also shown in the chart. The values were calculated based on data from SCOPE database, and it includes the scheduled arrival and departure times of all of the flights at Helsinki Airport.

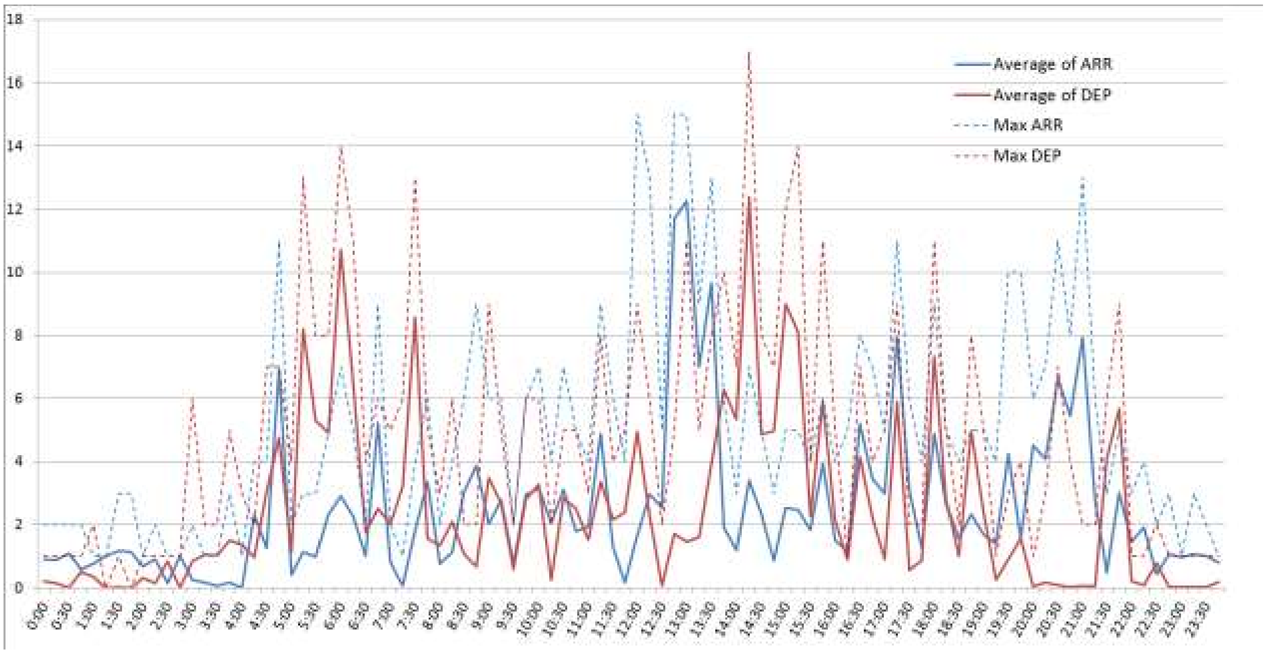


Figure 9: Average and maximum values of planned arrivals and departures per 15 minutes on the runway during a day in UTC time on the study period at Helsinki Airport (to convert to local time +2 hours need to be added)

In Figure 9 the blue lines are the planned arrivals and red line the planned departures. The morning bank shows first some arrivals following a larger peak of departures. The departing aircraft have been parking during the night at Helsinki Airport. The maximum of average and maximum values are 7 and 11 for arrivals and 11 and 14 for departures per 15 minutes. The afternoon bank has higher and wider peaks than the morning bank so this bank is more straining on the capacity of the airport. The values for the afternoon bank maximum of average and maximum values are around 12 and 14 for arrivals and 11 and 17 for departures per 15 minutes.

In Figure 10 the arrivals and departures are depicted with a bubble plot which visualizes the frequencies of different planned arrival/departure combinations in the study interval. The frequencies are in this case number of 15 minute intervals with the specified number of planned arrivals and departures. The study period between 1.11.2012 and 31.3.2013 constituted of 9680 15 minute time intervals with runway operations. The total time was 14400 15 minute time intervals in which 4720 didn't have any runway operations in them.

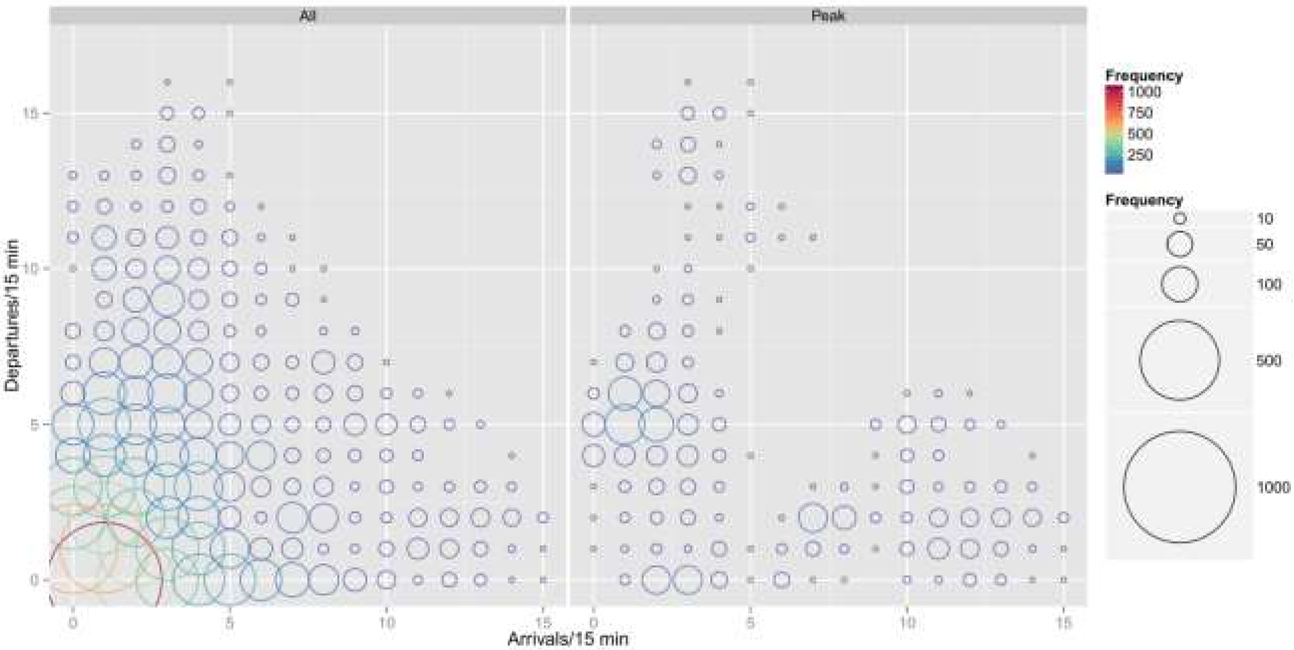


Figure 10: A bubble plot of all the arrival/departure combinations and their frequencies. The plots for all flights and the peak period flights are shown. The bubble sizes as well as its color indicate the frequency for the number of arrivals/departures per 15 minutes combinations. Study interval was between 1.11.2012 and 31.3.2013.

All of the points and their frequencies are depicted in the left plot and the afternoon bank is depicted in the right one of Figure 10. In the left plot it can be seen that the relatively low utilization (small amount of arrivals and departure) of runways are the most common. Over 7 planned arrivals or departures are pretty rare. That's because there are so many low operation times in the night and in the evening. In the right plot are the afternoon banks planned operations and their frequencies. Here the story is different. There are a lot of runway operation times with a lot of planned arrivals and a lot operation times with a lot of planned departures. The crooked structure of the planned runway operations is evident (this is also true for the realized runway operations). There are not a lot of times with a relatively high amount of both arrivals and departures.

To conclude this chapter, clearly the busiest time at the runway is the afternoon bank. In the following runway capacity analyses the estimated capacities are compared to the planned utilization, the planned runway operations on the afternoon bank described in this section.

4.2.3 Capacity and affecting factors

In this section a method for estimating the runway capacities in different situations is put into practice based on the theory represented in sections 2.3 and 3.1.1 and method described back in the beginning of section 4.2.1.

Some factors' effects on the capacity are studied in this section. The effects of factors on runway capacity studied are:

- Runway configuration
- Amount of fallen snow (in cm/h) during the same hour
- Visual conditions (*VFR*/*IFR* conditions)
- Wind, with categorized average wind speed conditions

These factors are included in the factors affecting runway capacity what were found in literature which were introduced in section 2.3. The factors were chosen because of some thoughts arisen based on interview and discussions with people at Finavia and Finnair, but also the availability of the weather data. Especially winter storms, when there is a lot of falling snow and possibly high winds and wind gusts, where identified as lowering the capacity of the runways.

The runway configuration affects the capacity at use at all times so it has to be fixed when trying to assess the effect of the different weather conditions on the capacity. In the following analyses the effects of other factors are assessed with keeping the runway configuration constant for each runway configuration separately.

4.2.3.1 Fitting capacity curve to the data

The procedure "rqss" from the Quantreg package is used for the piecewise linear quantile regression estimation. The lambda parameter introduces a penalty for the number of the linear curves fitted to the data. A Lambda parameter value of 0.01 is used in the estimation, which ensures that there are enough of linear curves in the estimation to embody the whole shape of capacity curve implicated by the data.

Restrictions on the shape of the capacity curve were added to the procedure. The estimation included restrictions that the curve should be

- Concave
- Decreasing

As the curve used to estimation was the 99.9%th quantile curve. This means that for each of the curves fitted only 1/1000 of the observations were rejected by the algorithm and considered as outlier in this capacity estimation.

The estimated capacity curves for the mostly used runway configurations are shown Figure 11. The most used runway configurations are:

- Arrivals on 22L / Departures on 22R
- Arrivals on 22L / Departures on 22L
- Arrivals on 15 / Departures on 22R
- Arrivals on 15 / Departures on 22L
- Arrivals on 04R / Departures on 04R
- Arrivals on 04L / Departures on 04R

The distribution of the used runway configurations in the study interval are shown for all of the configurations in Table 1. The rest of the used configurations constitute only 2,1% of all the measures and they have sample sizes of less than 100 so they are excluded from the study.

Table 1: Distribution of the runway configurations used in the study interval

Runway configuration	n	% of total
04L/04L	4	0,0 %
04L/04R	2075	21,4 %
04L/15	68	0,7 %
04L/22R	8	0,1 %
04R/04R	1470	15,2 %
04R/22R	16	0,2 %
15/04R	17	0,2 %
15/15	34	0,4 %
15/22L	132	1,4 %
15/22R	2674	27,6 %
22L/04R	16	0,2 %
22L/15	7	0,1 %
22L/22L	163	1,7 %
22L/22R	2958	30,6 %
33/22R	27	0,3 %
33/33	11	0,1 %
TOTAL	**9680**	**1**

The fitted maximum capacity curves at 99.9% quantiles for the 6 mostly used runway configurations different runway configurations are depicted in Figure 11. The frequencies of arrival/departure combination observations are depicted as 3D-histograms. The points near the origin (0,0) are much more frequent and have longer columns than the ones near the edge. Because the data consists of observations from all times of day, the columns near the fitted capacity curve are very low meaning that the near capacity utilization of the runways is very rare. Had the histograms been assessed for the peak period the highest columns would have been little more in the middle of the capacity area. Anyway the utilization of the runways is pretty low.

The capacity curves seem to fit very well to the data, leaving out only few outliers. As previously stated (section 4.2.1.2), the same runway arrival and departure configuration estimations are more prone to error when trying to assess their capacities. This is why the 22L/22L configuration and 04R/04R might be more unreliable than the other estimates. To conclude this method seems to be reasonably good in estimating the capacity curves at Helsinki-Vantaa airport.

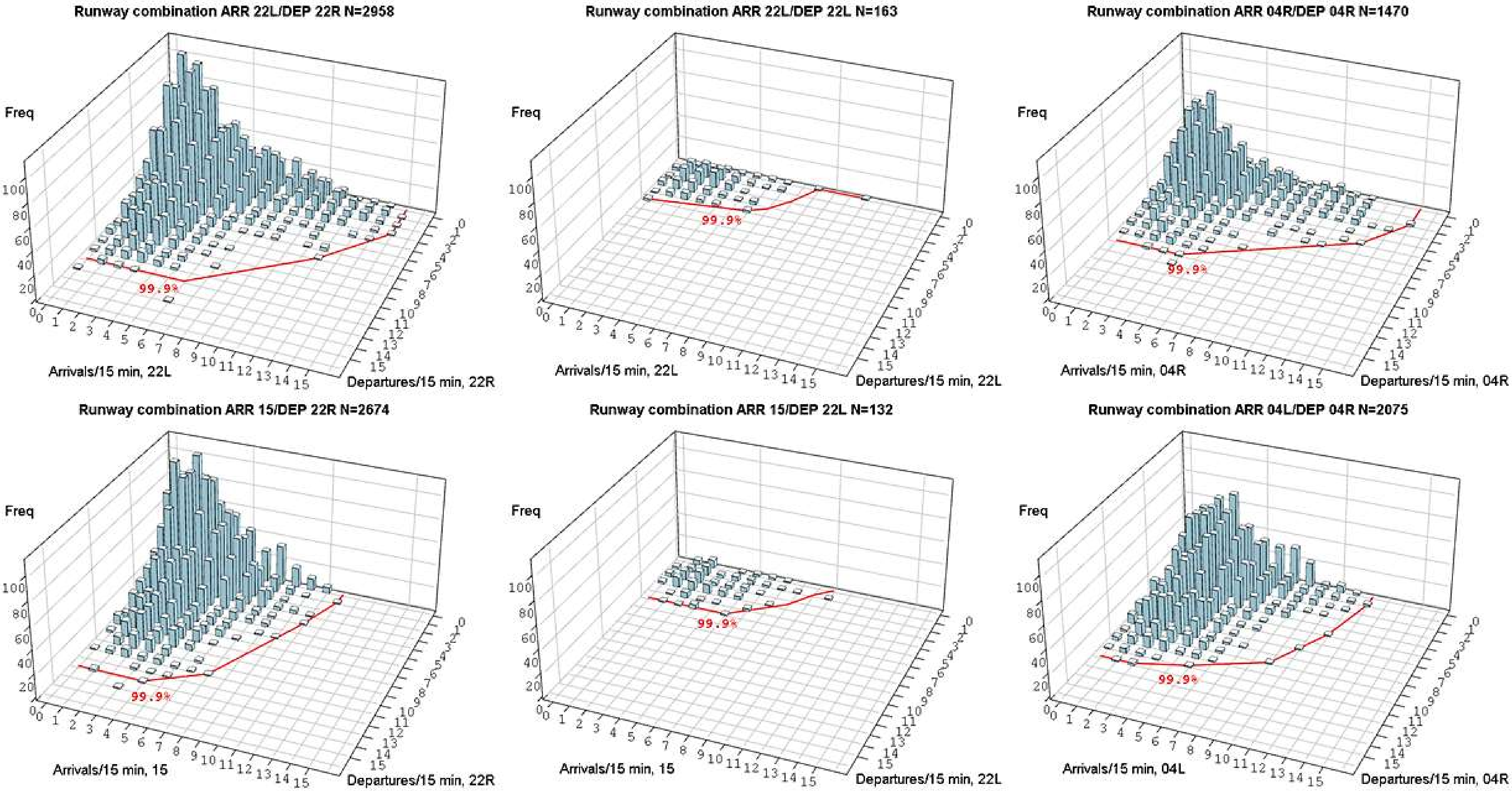


Figure 11: Estimated capacity envelopes for most used (n>100 15min time intervals) runway configurations

4.2.3.2 Effect of runway configuration on capacity

The runway configuration in use has a great impact on the capacity of the runways. The fitted maximum capacity curves for the different runway configurations are depicted in the Figure 12. The four mostly used capacity curves have a high sample sizes so they can be considered relatively good approximations of the capacity envelopes. The sample sizes for the runway configurations 04L/15, 15/22L and 22L/22L are pretty low so it's difficult to say how near these curves are to true maximum capacity when the configurations are in use. They can be considered at least a lowest boundary for the maximum capacities; the true capacities cannot be lower than the ones here.

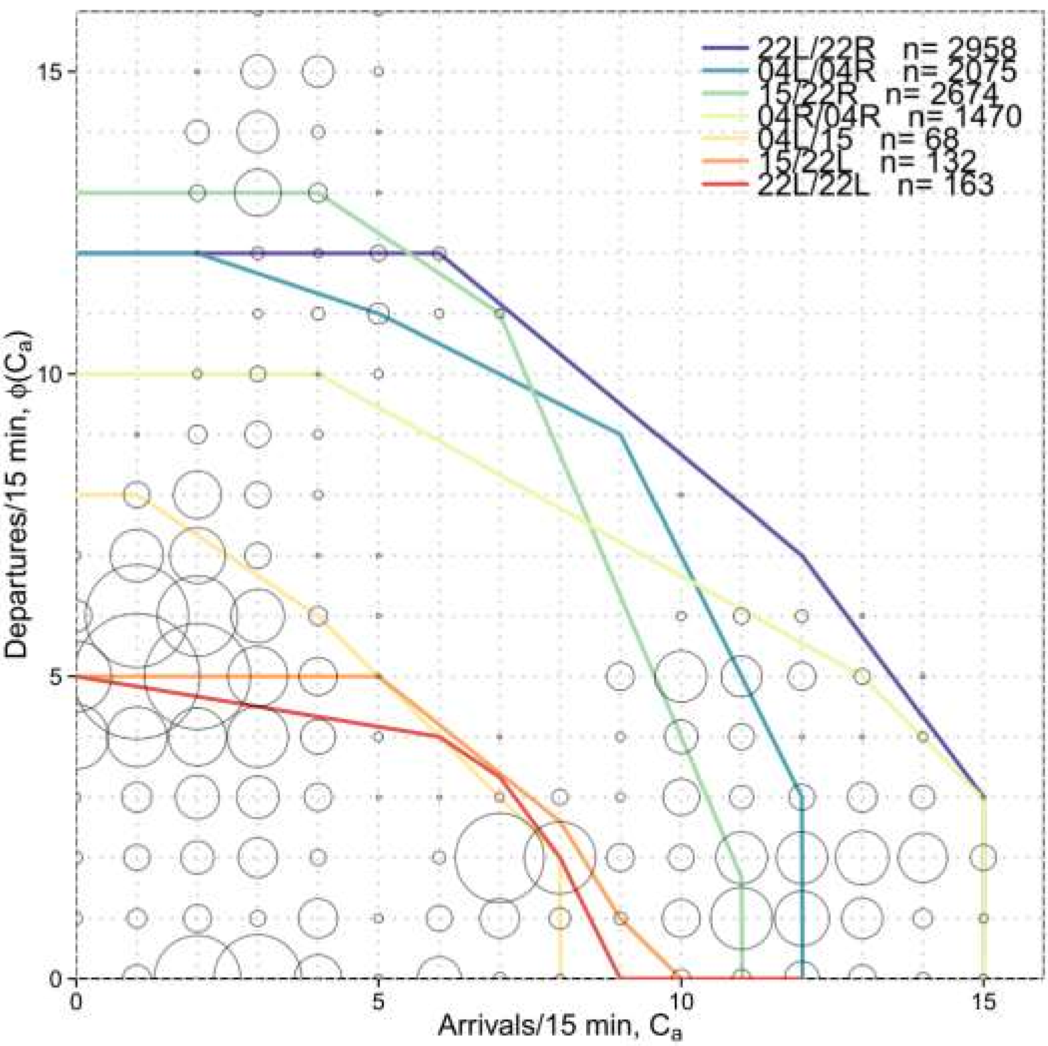


Figure 12: Estimated runway capacity curves for the mostly used runway configurations at Helsinki Airport

The largest capacity is with the configuration 22L/22R. The next ones are 04L/04R, 04R/04R and 15/22R. It's not possible to put these configurations in order because they might have either the arrival capacity or the departure capacity larger than the others. The 04R/04R configuration capacity curve comes with the previously presented flaws.

The clearly smaller capacity configurations have been 04L/15, 15/22L and 22L/22L, though they have small sample sizes. Because of the small amount of operations with these configurations the results aren't that reliable. Nevertheless the capacities of these configurations are much, much lower than the ones presented in the previous paragraph. They are even as low as only 5 departures or 8 arrivals per 15 minutes.

The Figure 12 also includes the planned arrivals and departures during the peak period of the day, which were introduced in Figure 10. Table 2 summarizes the number of 15 minute time intervals that would have been within the estimated capacity curves in a specific runway configuration. The figures give an estimate of how many of the planned operation times would not have been operated fully in the planned amount of operations. The percentage of planned operations that could have been operated is higher than the calculated amount, but would be difficult to calculate. This is why the %within capacity (of the planned peak period operations) gives more pessimistic view of the capacity and lower percentages than the percentage of feasibly operable planned operations, which was not calculated.

Table 2: The relative percentages of how many of the planned flight operations could be executed within capacity when a certain runway configuration is in use.

Runway configuration	Within capacity	Outside capacity	% within capacity
22L/22R	1703	99	94,5 %
04L/04R	1540	262	85,5 %
15/22R	1410	392	78,2 %
04R/04R	1673	129	92,8 %
04L/15	1148	654	63,7 %
15/22L	944	858	52,4 %
22L/22L	671	1131	37,2 %

A striking observation is that some of the planned runway operations are never feasible to be conducted. The results suggest that they could not be carried out in any of the runway configurations. There certainly wasn't any planned 15 minute times during the study interval that the runway has operated with as many departures has were planned (the points in outside all the capacity curves).

The configuration in use is chosen by the *ATC* based on the operating conditions: the flight schedule, air traffic and weather conditions. One of the most important feats when deciding the configuration is the wind direction and its velocity. In that sense the runway configuration already has some inherent operating conditions that affect the capacity when a specific configuration is used this is good to keep in mind when interpreting the results.

4.2.3.3 Effect of snowfall on capacity

The effect of falling snow was on the runway capacity is studied here. The falling snow and winter storms are often creating delays at Helsinki Airport during the winter. In Figure 13 are depicted the effects of snowfall on capacity with the four most used runway configurations in snowfall conditions. The envelopes were considered not statistically significant if there was less than 20 observation points, so in the configuration 22L/22R the heavy snowfall's effect was left out.

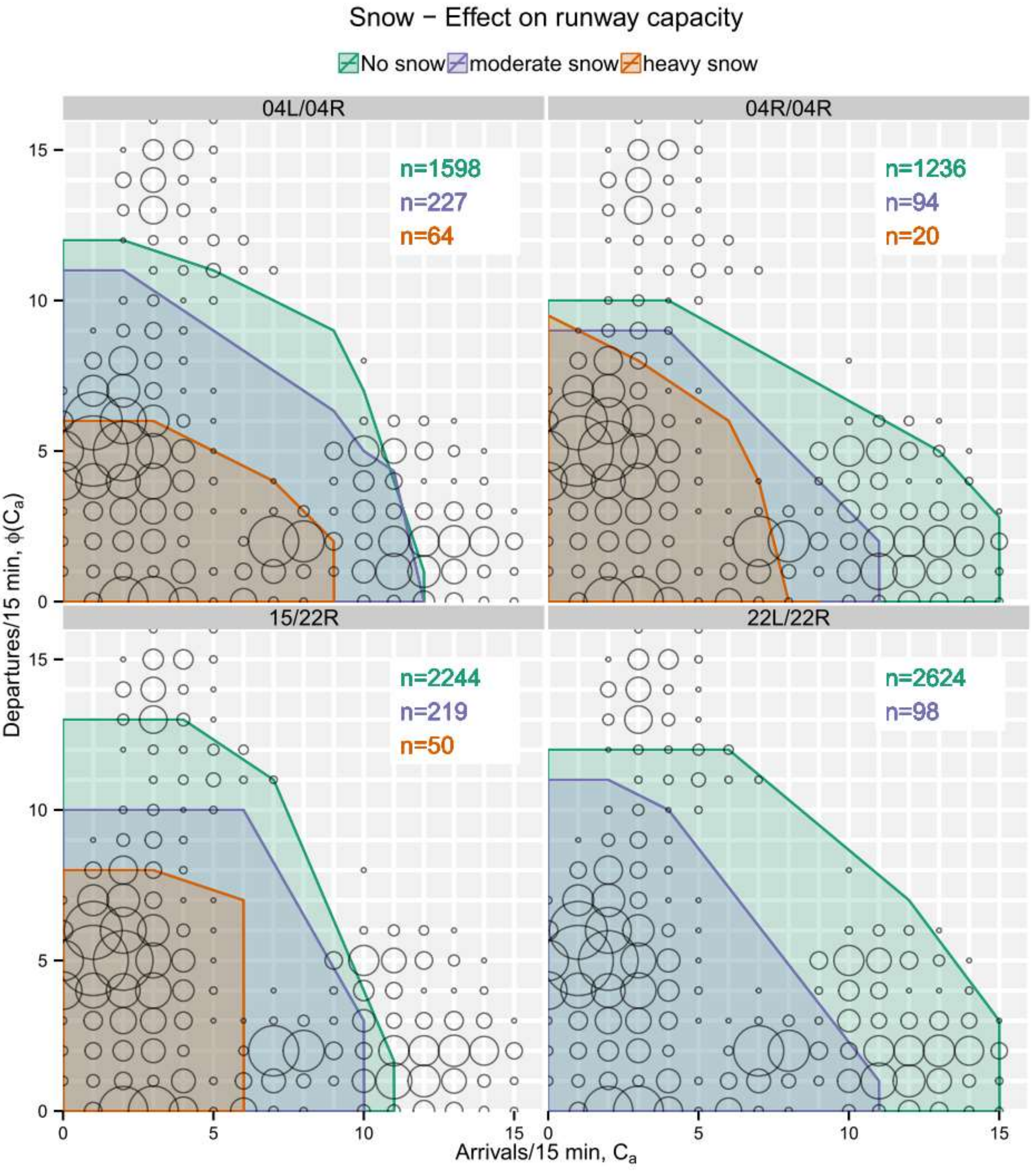


Figure 13: Effect of snow on runway capacity, factored by snow depth change during the same hour.

The snowfall amount was factorized to no snow, moderate snow and heavy snow conditions during the hour that the 15 min observation point took place. The factor no snow means that there was no snow or very little. The moderate snow factor means 1 cm change in the snow cover during an hour and heavy snow means 2 cm or more change during an hour.

The planned runway operations during the peak period from Figure 10 are also depicted in the diagrams of Figure 13. It helps to see how many of the planned runway operations are inside the estimated capacity curves.

The fitted capacity curves seem to indicate that the effect of heavy snowfall is quite drastic. The moderate snowfall has also a significant effect in all the runway configurations. It has to be considered that the sample sizes for the heavy snowfall are very low so it could be that the heavy snow capacity curves are little bit underestimated. The effects are clear and the shapes of the capacity curves follow the capacities of unaffected runway configurations pretty well.

The relative frequencies of snowfall categories in the study interval are presented in Table 3. Any significant snowfall happened 8,4% of time, which means snowfall affected relatively high proportion of runway operations.

Table 3: The relative frequencies of different snowfall conditions during the study interval 1.11.2012-31.3.2013

SnowLevel	Frequency	Total N	Relative frequency
No snow/light snow	8868	9680	0,916
moderate snow	659	9680	0,068
heavy snow	153	9680	0,016

To try to quantify the potential magnitude of the effect, all the planned runway operations were calculated whether or not they are within the capacity curve in each situation. The figures are presented in Table 4. This figure indicates on average how many of the planned runway operation times during the peak hour would have been affected by the snowfall, had to snowfall occurred at the specific time of operation.

The proportions of the runway operations within capacity are quite small. If the figures were counted for runway operations at all times of day they would of course be higher because the points with lower amount of operations are more frequent at times outside the peak hour and more of them would be within capacity curves even when the capacities are limited by the prevailing snowfall.

Table 4: The relative percentages of how many of the planned flight operations could be executed within capacity in different snowfall conditions.

Runway configuration	Snow level	Within capacity	Outside capacity	% within capacity
04L/04R	No snow	1505	297	83,5 %
04L/04R	moderate	1447	355	80,3 %
04L/04R	heavy	1140	662	63,3 %
04R/04R	No snow	1669	133	92,6 %
04R/04R	moderate	1355	447	75,2 %
04R/04R	heavy	1182	620	65,6 %
15/22R	No snow	1435	367	79,6 %
15/22R	moderate	1324	478	73,5 %
15/22R	heavy	1093	709	60,7 %
22L/22R	No snow	1701	101	94,4 %
22L/22R	moderate	1346	456	74,7 %

The percentages of planned operation points (during peak period) within capacity limits range from 94,4% for configuration 22L/22R in no snow conditions and to as low as 60,7% for runway configuration 15/22R in heavy snow conditions. The average effect of the moderate snowfall on planned peak runway operation times was 11,6% and for heavy snow 22,1%.

It's important to understand that the possible effects are larger if the planned operations at peak period increase. A need for further study would be beneficial to find out why the capacity is affected in the ways presented.

4.2.3.4 Effect of high winds on runway capacity

The effect of average wind speed on the runway capacity was studied. The wind conditions were factored in to two distinct categories of hourly average wind speed conditions: normal and high. The threshold level was set at level 8,4 m/s, below which the average wind speed was around 95% and above 5% of time. The capacity envelopes in the different runway configurations and wind conditions are depicted in Figure 14.

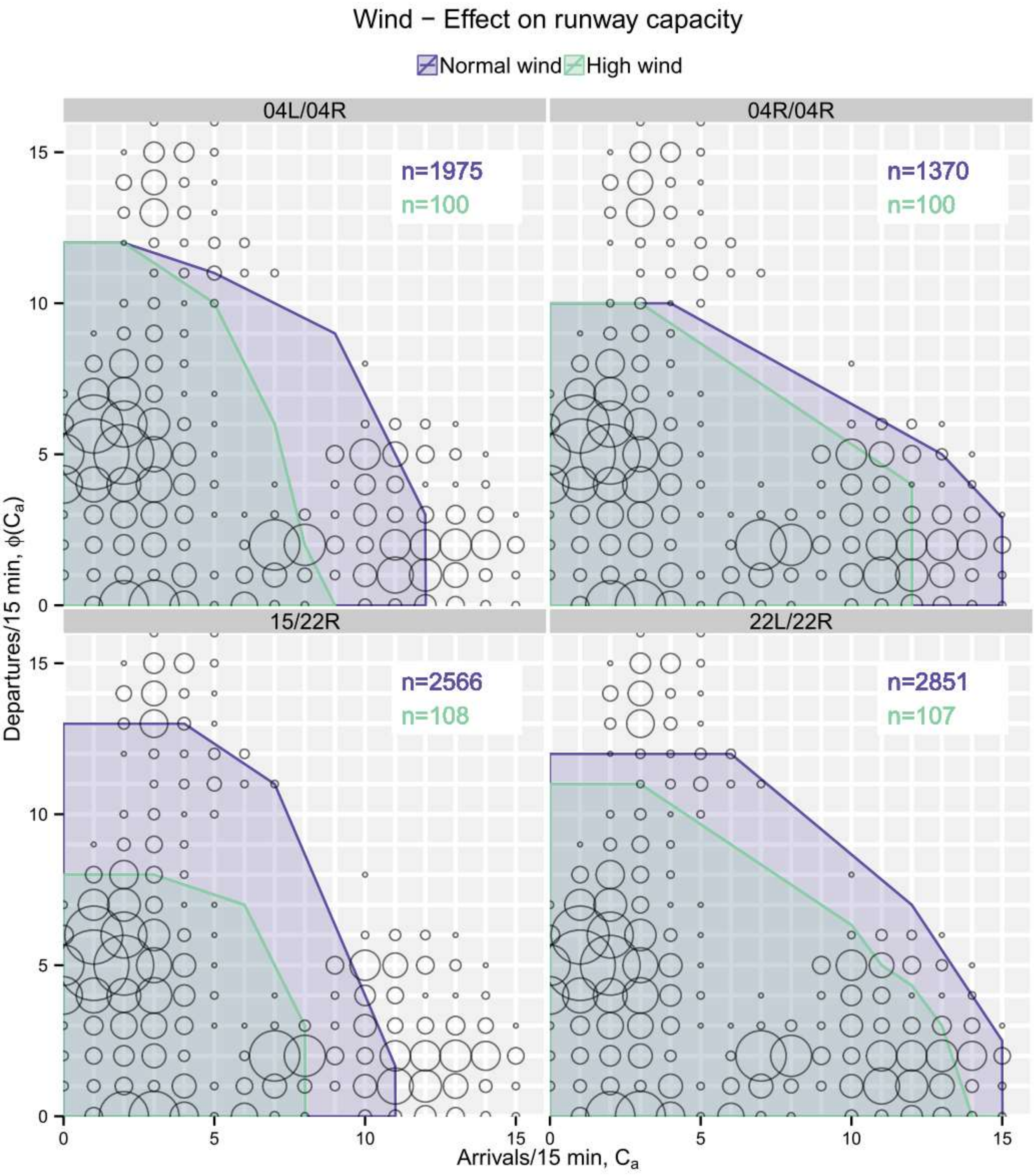


Figure 14: Effect of wind on runway capacity, factored by average wind speed in the same hour.

Above we see the planned runway operations during the peak periods depicted in the Figure 14. We perceive an odd change in the capacity curves in runway configurations 04L/04R and 04R/04R. It seems that only the number arriving aircraft has changed. In configuration the 22L/22R in the tails of the capacity seems to have reduced one operation and two operations in the middle of the capacity envelope. The greatest reduction was seen in the15/22R configuration where the high winds have reduced the runway operations by 3 to 6 operations in different parts of the capacity curve.

The effect of wind on capacity is not as clear as the effect of snowfall. For example there seems to no effect on the capacity on the left side of the curves in runway configurations 04L/04R and 04R/04R. On the other hand the other two runway configurations have clear effects throughout the curves. The effects seem to vary much from configuration to configuration. This puts some doubt whether the effect has been caused by some covariate factors that happen simultaneously with high average wind speeds.

The relative frequencies of wind categories in the study interval are presented in Table 5. The high wind in this setting affected the capacity 4,6% of time and the effect was not that clear.

Table 5: The relative frequencies of normal and high wind conditions during the study interval 1.11.2012-31.3.2013

Wind level	Frequency	Total N	Relative frequency
Normal wind	9234	9680	0,954
High wind	446	9680	0,046

In Table 6 are the relative percentage of planned operating times that would have been affected, when they would have been restricted by the high wind capacity curves. The average reduction, in the planned operations within capacity, of the high wind effect was 10,8%.

Table 6: The relative percentages of how many of the planned flight operations could be executed within capacity in different wind conditions.

Runway configuration	Wind level	Within capacity	Outside capacity	% within capacity
04L/04R	Normal	1546	256	85,8 %
04L/04R	High	1265	537	70,2 %
04R/04R	Normal	1673	129	92,8 %
04R/04R	High	1522	280	84,5 %
15/22R	Normal	1440	362	79,9 %
15/22R	High	1235	567	68,5 %
22L/22R	Normal	1709	93	94,8 %
22L/22R	High	1607	195	89,2 %

4.2.3.5 Effect of Visibility on capacity

The effect of visibility on runway capacity was also studied. In their article Ramanujam and Balakrishnan, 2009 studied also the effects visual flight rules and instrument flight rules on the capacity in a very similar manner as in this thesis. There was no data to indicate whether a single runway operation was really conducted under *VFR* or *IFR* regulation in the data. A factorized variable of horizontal ground visibility with two levels is introduced to account for the two rules. In the standard of *ICAO* (Convention on International Civil Aviation – Rules of the Air), which is also followed in Finland, one of the visual flight rules regulation is that there has to a minimum of 5 km of horizontal ground visibility. Normal visibility level is set to simulate the *VFR* rules bad visibility factor is set to simulate the conditions in *IFR* flight rules. The threshold value was set below the standard to 3km horizontal visibility, to simulate the *IFR* conditions. There are also other regulations in the standard that require the *IFR* rules, but weather data on the relevant conditions were not available.

The Figure 15 summarizes the capacity curves for the different configurations and visibility levels. Here too the planned runway operations are depicted in the figure to visually see the operating points that would not be possible to conduct at the runway in different capacity situations. The capacity curves in bad visibility follow very closely that shape of the unrestricted capacities of the runway configurations, but are reduced.

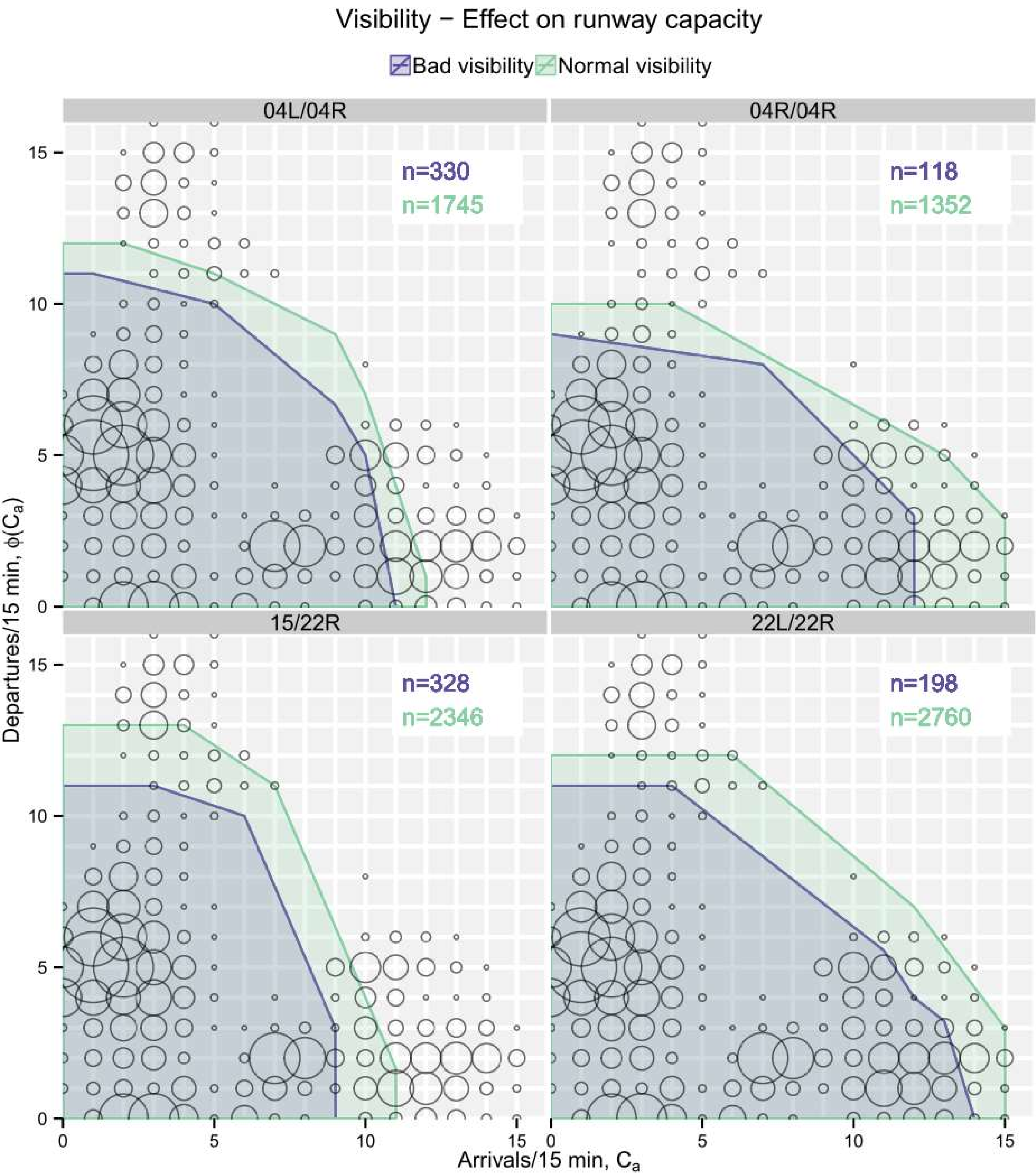


Figure 15: Effect of bad visibility on runway capacity, factored by horizontal ground visibility in the same hour.

The effect of the bad visibility varies mainly from a reduction of 1 to 2 runway operations per 15 minutes. The results are probably pretty reliable, because of the closely followed shapes of the capacity envelopes. The estimated capacity envelopes are probably pretty valid as well because of the relatively high sample sizes (over 100 in each) in the estimates.

In Table 7 are the relative frequencies of the visibility levels. The bad visibility i.e. under 3km horizontal ground visibility was observed 11% of time of the operational times at runways during the study interval.

Table 7: The relative frequencies of bad and normal visibility conditions during the study interval 1.11.2012-31.3.2013

Visibility level	Frequency	Total N	Relative frequency
Bad visibility	1062	9680	0,110
Normal visibility	8618	9680	0,890

The percentages of how many times of the planned runway operation points could not have been conducted as planned, had the capacity restriction affected all of the points are in the Table 8. The average decrease in the percent of points within capacity in bad visibility is 7,8%, so the effect is a lot smaller to the current planned number of flights than in the previously covered effects of snowfall and wind.

Table 8: The relative percentages of how many of the planned flight operations could be executed within capacity in different visibility conditions.

Runway configuration	Visibility	Within capacity	Outside capacity	% within capacity
04L/04R	Normal	1505	297	83,5 %
04L/04R	Bad	1370	432	76,0 %
04R/04R	Normal	1673	129	92,8 %
04R/04R	Bad	1499	303	83,2 %
15/22R	Normal	1440	362	79,9 %
15/22R	Bad	1282	520	71,1 %
22L/22R	Normal	1710	92	94,9 %
22L/22R	Bad	1612	190	89,5 %

4.2.4 Error sources in estimation of the factor effects

One of error sources the data represented only on winter and no summer data was analyzed since the study interval consisted of data between 1.11.2012 and 31.3.2013. The operating conditions could be different in summer affecting the capacities and the results.

Helsinki Airport operates so rarely (very short time a day) at near the capacity and the runway operations are rarely congested, which affects the reliability of the results. To more reliably estimate the capacities the runways would need to be more often in the state of congestion, when there would be more near capacity data available. There is no standard way of assessing the reliability and validity of the results. The results are probably quite reliable, but because of the quite small sample sizes in some effects they are not valid meaning here that the capacity curves aren't necessarily completely at the right place.

The reasons for the capacity curve reduction aren't also straight forward. There might be underlying reasons for the capacity reduction might differ from the factors presented. It's not possible to completely isolate the effect and there might be some covariates that affect the capacities. The studying of these effects is out of scope of this thesis.

The effects of the factors might also be continuous or have many thresholds so the capacity curves could be estimated for all levels of factors. In this study the factors effects were estimated only on certain levels that were chosen arbitrarily.

The effects of combined factors were not assessed. If two or more factors are affecting the runway operations at a time there would be less capacity available. In a specific situation the capacity might then be lower than what is presented by so the effect of some unstudied or studied covariate.

4.3 Deicing process capacity

Deicing is the process of treating the aircraft with a liquid substance that removes ice by lowering the freezing point of water and prevents ice from forming on the aircraft. The de-icing processing is important before taking off in the winter, because ice on the aircrafts body affects the aircraft's aerodynamic properties. Taking off with an aircraft with ice on its body is very risky. This is why in winter in certain weather conditions this treatment has to be done,

There are three deicing processing locations at Helsinki-Vantaa. Those are remote 6, remote 8 and apron. The apron deicing processing took place at the gates 25-30. The locations of the deicing processing stations in winter 2012-2013 are shown in Figure 16. The remote 6 deicing location is used for deicing narrow body planes taking off to the direction 220°. The remote 8 on the other hand is used for deicing narrow body taking off to the direction 40°. The apron is used to conducting deicing treatment for both narrow and wide bodies.

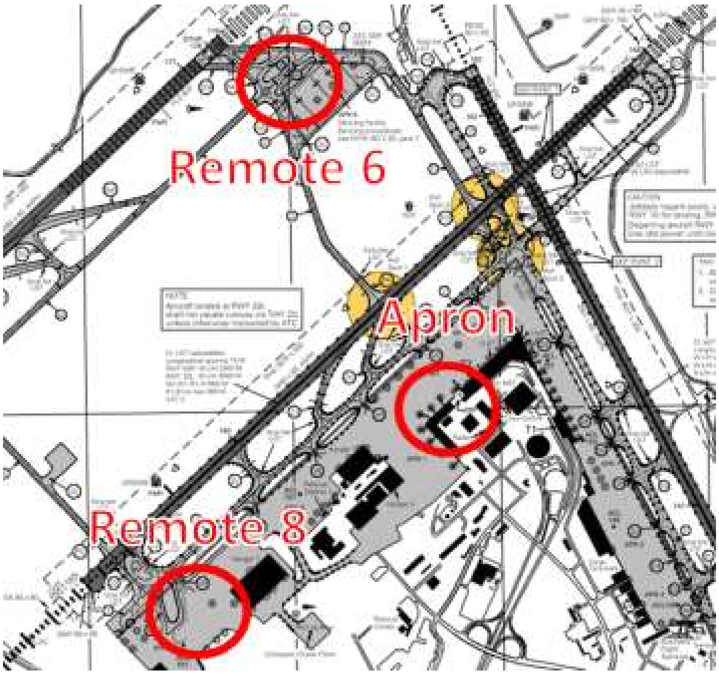


Figure 16: Deicing locations Helsinki Airport (Finavia Oyj, 2013b)

The output of the deicing process is the deicing treatment done (the provided service) to the plane. Its capacity is measured as treatments done per unit of time. It is clear that the time taken to process on plane is inversely proportional to the number of planes that can be treated in an hour in a single processing line. So the expected Deicing capacity C_{DI}, can be expressed as,

$$C_{DI} = \frac{1}{E(t)} [treaments/\text{hour}] \quad (1)$$

where $E(t)$ is the expected time taken to process a single plane in hours.

There are different conditions or factors that affect the treatment time of single plane. Let A, B, C be the set of all the different combinations of conditions that have known effect on the treatment time.

$$A, B, C, \ldots \text{are sets of categorized conditions for a deicing treatment}$$

The set A can be for example the types of aircraft or categorized average wind speeds. K is a set all the possible combinations of the discreetly categorized condition sets,

$$K = \{(k_A, k_B, k_C \ldots) \mid k_A \in A\ k_B \in B, k_C \in C \ldots\ \}$$

I is an index set for all the elements of K

$$\left\{ i \in I \mid K = \bigcup_{i \in I} K_i \right\}$$

If the conditions can be forecasted for the next day and weights w_i can be assigned to each condition combination $\in I$. Then w_i are the weights of proportionally how many aircraft are treated in conditions $i \in I$ (i.e. w_i is how many planes are to be treated in conditions i when divided by the total number of planes to be treated).The theoretical total capacity of deicing treatments for a single processing line can be calculated with formula (if it can be assumed that the next treatment can start straight away after the one before):

$$C_{DI} = \frac{1}{\sum_i w_i E(t_i)}, \qquad \sum w_i = 1, \qquad (2)$$

The different processing lines can be marked with $j \in \{1..N\}$, where N is the number of processing lines. When there are more than one processing line and the weights w_{ji} for different conditions K_i are different for the processing lines j, the expected total capacity C_{DI} can be expressed as:

$$C_{DI} = \sum_j \frac{1}{\sum_i w_{ji} E(t_i)}, \ \ \sum w_i = 1 \qquad (3)$$

In order to be able to calculate the total capacity the expected treatment times t_i need to be estimated for the set of condition combination $K_i,\ i \in I$. In the next section a model for estimating the expected value of deicing times is introduced.

4.3.1 Deicing processing time model

To create a theoretical model for calculating the deicing capacity, the expected deicing time t_i needs to be estimated for the different conditions K_i. In section 3.1.2 the statistical model gamma regression to estimate the expected deicing time is introduced. Gamma regression should be considered when the response variable is non-negative and has the possibility of having large positive values from zero to unlimited ($[0,\infty[$).

4.3.1.1 Data

The data was gathered from the CAPCO database of Finavia, which has all the deicing process data related to the deicing processing taking place in the winter at Helsinki Airport. The data had in the time interval 12000 rows of data of which only around 1700 where chosen to be included in the study. Only the rows which had both the deicing starting time and ending time reliably reported were used in the deicing time estimation. Large number of deicing times processed from the data didn't have any ending time reported and the data had also other inconsistencies as well. The troubles with the data might be reflected in the results, but their effects are hard to assess.

4.3.1.2 Model

A histogram of valid deicing times at Helsinki-Vantaa and a gamma distribution fitted to the deicing data is depicted in Figure 17. It can be perceived from the figure that the distribution fits somewhat well to the deicing duration data. The p-value for the Kolmogorov-Smirnow test statistic for this maximum likelihood fitted distribution is 0.05001308. On 10% significance level we can say that Kolmogorov-Smirnow test doesn't reject the null hypothesis that this sample is from Gamma distribution. This would indicate that the gamma regression could be a good choice to model the relationship between different factors and the deicing processing time.

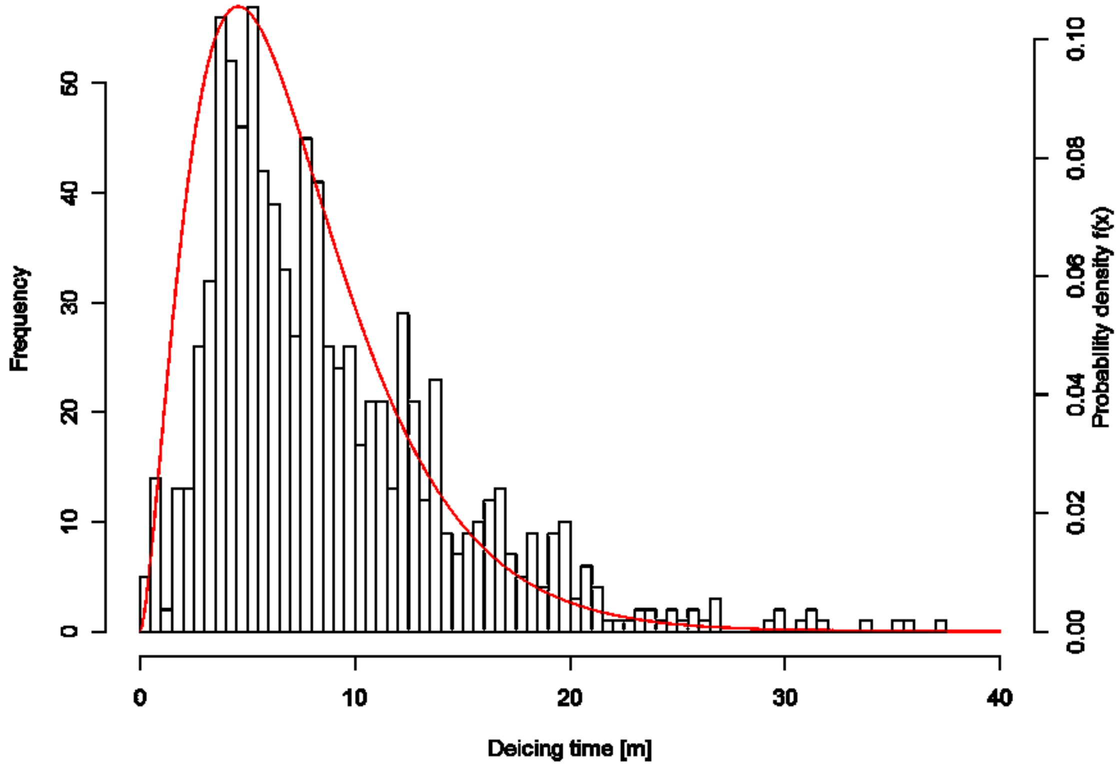


Figure 17: Deicing times from winter 2012-2013 and a fitted Gamma distribution

The effects of the following factors were studied to the expected deicing time and their coding in the following formulas:

- The deicing location (DL)
- The aircraft type (Narrow body/ Wide body) (AC)
- Number of deicing trucks conducting the treatment (TR)
- Amount of snowfall (SF)
- Temperature (TE)
- Wind velocity (WI)
- Relative humidity of the air (HU)

The purpose of the model is to estimate the expected value deicing times when affected by different factors. The expected mean for the deicing time t of the gamma regression model is of the form:

$$E(t) = \mu_i = \beta_0 + \beta_1 x_{DL,1} + \beta_2 x_{DL,2} + \beta_3 x_{AC,1} + \beta_4 x_{TR,2} + \beta_5 x_{TR,3} + \beta_6 x_{TR,4} + \beta_7 x_{SF,1} + \beta_8 x_{TE,1} + \beta_9 x_{TE,2} + \beta_{10} x_{WI,1} + \beta_{11} x_{WI,2} + \beta_{12} x_{HU,1} + \beta_{13} x_{HU,2} \quad (4)$$

All explanatory variables are factor variables except the amount of snowfall during two latest hours before the treatment $x_{SF,1}$. It takes whole numbers between 0 and 5, in which interval the model is fitted.

$$x_{SF,1} \in \{0,1,2,3,4,5\}$$

All the factorized explanatory variables are either 0 or 1,

$$x_{lm} = 0 \wedge x_{lm} = 1 \quad \forall \; l \in \{DL, AC, TR, TE, WI, HU\}\,, m \in \{1,2,3, \ldots\}$$

, where l is the index of the name of the categorized factor and m is the index of its level. Only one level of factorized variable for one factor is effective at a time which means that only one or none of the indicator variables x_{lm} variables for factors l can be 1 at a time,

$$\sum_m x_{lm} \leq 1\,, l \in \{DI, AC, TR, TE, WI, HU\}, m \in \{1,2,3, \ldots\}$$

The effects on the mean of the deicing times are affected by the beta coefficients (in minutes),

$$\beta_0 = The\ intercept\ term\text{:}\ Deicing\ location\ apron,\ Narrow\ body,$$
$$No\ snowfall,\ Temperature\ above\ zero,\ \ Normal\ wind,\ Normal\ humidity$$

$$\beta_1 = DeIcing\ location\text{:} apron \rightarrow remote\ 6$$

$$\beta_2 = DeIcing\ location : apron \rightarrow remote\ 8$$

$$\beta_3 = Narrow\ body \rightarrow \ Wide\ body$$

$$\beta_4 = TruckNo2 : 1truck \rightarrow 2\ trucks$$

$$\beta_5 = TruckNo3 : 1truck \rightarrow 3\ trucks$$

$$\beta_6 = TruckNo4or5 : 1truck \rightarrow 4\ or\ 5\ trucks$$

$\beta_7 = Centimeters\ of\ snowfall\ in\ same\ and\ previous\ hour$

$\beta_8 = Temperature\ above\ zero \rightarrow Temperature\ sub\ zero$

$\beta_9 = Temperature\ above\ zero \rightarrow Temperature\ below - 7°C$

$\beta_{10} = Normal\ wind \rightarrow Wind$ between $5\mathrm{m/s}$ and $8\mathrm{m/s}$

$\beta_{11} = Normal\ wind \rightarrow Wind\ above\ 8m/s$

$\beta_{12} = Normal\ humidity(< 80\%) \rightarrow High\ humidity(> 80\%)$

$\beta_{13} = High\ humidity(> 80\%) \rightarrow Very\ high\ humidity\ (93\% - 100\%)$

The chosen fitted model and the estimated values of the beta β coefficients are represented in an output of the R glm procedure in Figure 18. The p-value for a chi-squared test statistics with null hypothesis that the expected value of the model is of the presented form is 1.552354e-33, which doesn't reject the model for the expected value.

All of the coefficients except the TruckNoFactor2 and TemperatureSubZero are statistically significant in 1% level of significance. No model was found were the TruckNo2 factor level would be made statistically significant. It can be intuitively argued that there should be significant change in the deicing time for instance when conducting the treatment with 1 or 2 trucks but the model didn't incorporate that. Possible reasons could be that there are problems with the data or that some very significant factor is missing from the model. The statistical insignificance of subzero temperature could on the other hand be due to the nonlinear relationship of the temperature on the deicing time.

The intercept term is 5.8 minutes which is the starting point for the expected value of the deicing time. When deicing is conducted at remote 6 or remote 8 location the effect on the expected value is a reduction of around 2,5 minutes for both deicing locations. The result is expected because the deicing can be done more quickly at these locations. The effect of the number of trucks used in the deicing treatment is around 0,44 minute reduction when done with 2 trucks, around 1,5 minutes with 3 trucks and around 1,8 minutes with 4 or 5 trucks. The cumulative snow depth change in the previous and current hour amount had an effect of 0,4 minutes per 1 centimeter of snow. Subzero temperature had an effect of around 0,4 minute increase and the freezing temperature (<7°C) had an increased amount of 0,9 minutes. For the wind effect, the medium wind conditions (5-8m/s) increases the deicing time for 1,44 minutes and high wind conditions (>8 m/s). High humidity (80-93%) increases the expected time for 1,5 minutes and very high humidity (93-100%) increases the expected deicing time for 2,7 minutes.

Altogether the effects of the different coefficients are logical, have the right sign and seem to be in accordance with what is to be logically expected. For example the deicing treatment is faster at the remote 6 and remote 8 where the treatment is more controlled and the effects on the de-icing time are roughly the same at both locations. Treatment of wide body aircraft takes longer according to the model as expected and increasing the number of trucks involved in the de-icing processing reduces the time taken by the de-icing. Lower temperatures make the de-icing time longer as well as higher average wind speed. Humidity also affects by increasing the deicing time the higher the relative humidity of the air is.

The Nagelkerke pseudo R-squared for the model is 0.307 which means that the fitted model has a lot of variation that's left unexplained by the model. This could indicate that there might be some other factors that should be included in the model in order to better predict the deicing time. It could be also due to some other error sources that are discussed in section 4.3.1.4 Error sources. The large variance could also be due to the fact the process actually is as variable as the model suggest.

```
Call:
glm(formula = formula, family = Gamma(link = identity), data = deicing
    subset = outlierLogicalVector, start = coef(fitLin))

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-2.7184  -0.4340  -0.1164   0.2148   2.0391

Coefficients:
                            Estimate Std. Error t value Pr(>|t|)
(Intercept)                   5.7368     0.4247  13.509  < 2e-16 ***
DEICINGAREAFactorRemote 6    -2.5497     0.2473 -10.312  < 2e-16 ***
DEICINGAREAFactorRemote 8    -2.5488     0.2622  -9.721  < 2e-16 ***
ACCategoryWideBody            1.7576     0.3371   5.213 2.08e-07 ***
TruckNoFactor2               -0.4478     0.3609  -1.241  0.21479
TruckNoFactor3               -1.5423     0.3932  -3.922 9.11e-05 ***
TruckNoFactor4 or 5          -1.8102     0.3638  -4.976 7.13e-07 ***
SnowCumul2hPosit              0.4373     0.1530   2.858  0.00431 **
TemperatureFSubZero           0.4295     0.2681   1.602  0.10934
TemperatureFFreezing          0.9141     0.3113   2.936  0.00337 **
WindFMediumWind               1.4381     0.1906   7.545 7.24e-14 ***
WindFHighWind                 3.6413     0.4417   8.243 3.27e-16 ***
HumidityFHigh humidity        1.4774     0.2421   6.101 1.29e-09 ***
HumidityFVery High humidity   2.6687     0.2661  10.030  < 2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for Gamma family taken to be 0.3187546)

    Null deviance: 713.12  on 1753  degrees of freedom
Residual deviance: 523.15  on 1740  degrees of freedom
AIC: 9277.7

Number of Fisher Scoring iterations: 8
```

Figure 18: Summary of the fitted gamma regression model on expected deicing time (R output)

Also other versions of the gamma regression models were considered. Namely, models where the continuous variables were not categorized in bins (factorized variables) and with different link functions.

The effects of meteorological conditions cannot be expected to be linear in nature. The meteorological variables are continuous in nature so they were categorized for the sake of clarity. In principle the categorization of variables in regression is not recommended, but it can be considered for instance when the effects are not linear. For example it cannot be expected that as the temperature rises above zero degrees centigrade that the deicing time would decrease at all. Also the effect of wind on the deicing time starts to kick in after 5m/s.

The categorization of the explanatory variables was done by using a mix of common sense and a regression tree models which suggested that certain threshold values for the meteorological variables were significant in terms of their effect on the deicing time.

Other link functions tried were the log link and the canonical inverse link. The results with these link functions didn't differ significantly from the identity link which was used in the model introduced here. Because the identity link provided the most intuitive and readily interpretable results it was chosen as the presented model here.

4.3.1.3 Model goodness and regression diagnostics

In Figure 19 top left corner are the deviance residuals of the fitted deicing regression model versus the predicted values. The average of the deviance residuals is around 0 with all predicted expected values of the deicing time which is how it is supposed to be. The residuals are distributed quite evenly well in between the [-2,+2] interval where should be around 95% of the values. There are a few very low values of nearly -3 on the deviance residuals, but over all there seems to be not obvious problems in this plot. In the scale-location the square root of the absolute values of the standardized deviance residuals are plotted against the fitted values of the regression model. Also there the residuals seem to be quite evenly distributed and the average value of the plotted values is quite constant. These plots indicate that there is no heteroscedasticity in the deviance residuals.

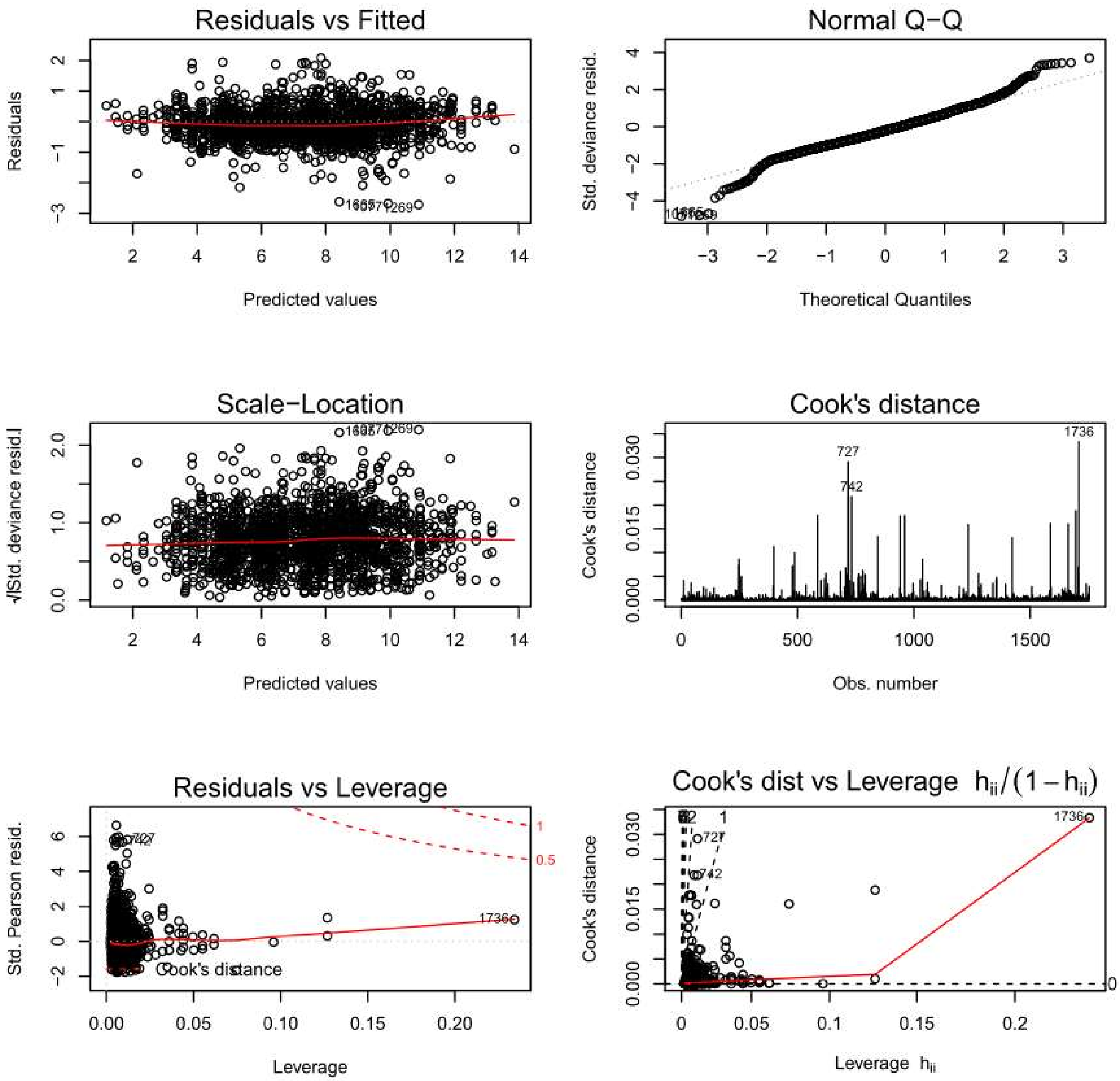


Figure 19: Residual diagnostic plots for the deicing time gamma regression model

In the top right corner is the Normal Q-Q plot where the quantiles of the distribution of the standardized deviance residuals plotted against the normal distributions theoretical quantiles. The deviance residuals should be approximately normally distributed when the model is fits the data well. When the model fits well to the data there should be approximately a straight line. It can be seen that in the left tail the in the Q-Q plot the quantile points fall below the straight line. This means that in the deviance residuals the high negative deviations from the mean are more common than in the expected theoretical distribution (in this case gamma distribution). In the right tail of the Q-Q plot the quantile points fall above the straight line, which means that in the deviance residuals the high positive deviations are also more common than in the expected theoretical distribution. In the middle the Q-Q plot follows the straight line very well. The Q-Q plot indicates that the both the low and high tails of the observed deviance residuals are little bit fatter than is to be expected by the model. Overall the Q-Q plot indicates the deviance residuals are very much like what is to be expected by the fitted model.

In the Cook's distance plot three observations were found to be having a little bit greater impact on the model parameter estimation than is to be wished. There are no other alarming observations found in the residual vs. leverage and Cook's distance vs. Leverage plots. The observations with higher Cooks distances were still found to have only a low impact on the regression coefficients (maximum of one hundredth of a minute) so no need for their censoring from the estimation was found.

Overall the fitted model seems to be much better than with standard linear regression. The model seems to be statistically significant apart from few factor levels of factorized variable in the model. The large variance in the model could indicate that there really is large variation around the expected deicing times or there are still some important factors that have been left out of the model or some other error sources causing the large variance.

4.3.1.4 Error sources

The data had a lot of missing treatment end times. That is why a lot of observations had to be discarded. It also seemed that the times are probably inputted in a way that causes much variance already because different people report them differently. Also many times all the trucks taken part in the treatment don't do treatment at the same length of time, but some of the trucks are idle or moving to next aircraft which affects the estimated model.

There could also operational issues that contribute to the inaccuracy or the variance of the model. For instance when there is congestion or urgency the deicing operators could be operating faster pace. On the other hand when there are more minutes left when the aircraft needs to depart the treatment might be slower because there is no need to work faster. This type of working slower and working faster will affect a lot how the parameters of the model are estimated.

4.3.2 Deicing capacity estimation

Formula for calculating the expected deicing capacity based on the presented regression model by combining the equations (3) and (4) is:

$$C_{DI} = \sum_j \left(\sum_i w_{ji} (\beta_0 + \beta_1 x_{DI,1} + \beta_2 x_{DI,2} + \beta_3 x_{AC,1} + \beta_4 x_{TR,2} + \beta_5 x_{TR,3} + \beta_6 x_{TR,4} + \beta_7 x_{SF,1} + \beta_8 x_{TE,1} + \beta_9 x_{TE,2} + \beta_{10} x_{WI,1} + \beta_{11} x_{WI,2} + \beta_{12} x_{HU,1} + \beta_{13} x_{HU,2}) \right)^{-1} \quad (5)$$

Where the weighted deicing times are summed over the processing lines and deicing locations. The factor indicator variables x_{lm} are set to values 0 or 1 according to which are the prevailing conditions K_i, expect for the $x_{SF,1}$ which represents the amount of snowfall in centimeters per two hours. This calculation of the capacity doesn't take into account the change times from one aircraft to another. That's why the capacity values calculated this way only reflect expected theoretical capacity and give an upper bound to the expected capacity. If the model included the change times (i.e. it would take into account the time it takes to change from one aircraft to another), the figures could be more reliable.

In Table 9 are presented expected capacities calculated with equation (5) and the regression coefficients from the model presented in Figure 18. The calculations were done in Excel and they required the solving of a small mixed integer linear programming model (MILP) with Excel's solver. The expected capacities are calculated for different number of apron and remote processing lines. It was assumed that all the "processing lines" have 2 trucks, which means that all the deicing treatments are expected to be conducted by 2 trucks. The capacities are also presented for three different weather conditions: easy, medium and hard. The details of the weather conditions used are:

- The easy weather conditions:
    - No snowfall

- Above zero temperature
- Less than 5 m/s average wind speed
- Less than 80% relative humidity.

- The medium weather conditions
  - 1 cm of snowfall during the latest 2 hours
  - Temperature between 0°C and -7°C
  - 5 m/s - 8m/s average wind speed
  - 80%-93% relative humidity.
- The hard weather conditions:
  - 2 cm of snowfall during the latest 2 hours
  - Temperature below -7°C
  - Over 8m/s average wind speed
  - Over 93% relative humidity.

Table 9: Calculated expected theoretical de-icing capacities based on the expected deicing time model.

Fleet mix (narrow/ wide)	Apron processing lines	Remote processing lines	Trucks per processing line	Number of trucks	Capacity/hour (expected theoretical)		
					Narrow body	Wide body	Total AC / hour
					Easy weather conditions		
3,3	1	1	2	4	23,7	7,2	**30,9**
3,3	1	2	2	6	43,8	8,5	**52,3**
3,3	2	2	2	8	47,4	14,4	**61,7**
3,3	2	3	2	10	65,7	17,0	**82,7**
					Medium weather conditions		
3,3	1	1	2	4	11,6	3,5	**15,1**
3,3	1	2	2	6	18,4	5,5	**23,9**
3,3	2	2	2	8	23,2	7,0	**30,3**
3,3	2	3	2	10	30,0	9,1	**39,1**
					Hard weather conditions		
3,3	1	1	2	4	7,5	2,3	**9,7**
3,3	1	2	2	6	11,6	3,5	**15,1**
3,3	2	2	2	8	14,9	4,5	**19,4**
3,3	2	3	2	10	19,0	5,8	**24,8**

The calculated expected capacities in *AC*/hour range from 9 to 82 aircraft per hour. The capacity figures can be compared with the histogram in Figure 20. The histogram includes the frequencies of number of departing aircraft per hour. The number of departing narrow body and wide body aircraft could not be distinguished from the data so the total number of departing aircraft needs to be used in the comparison. The departures per hour, between 1.11.2012 and 31.3.2013 ranged from 1 to 41 departures per hour. The numbers in Table 9 would suggest that with 10 trucks in two apron processing lines and 3 remote processing lines would in most cases be enough to care for the demand. Only in the case of the presented hard weather conditions the model would forecast that the 10 trucks could not handle all the demanded treatments of the departing aircraft. As discussed earlier the model doesn't take into account the changing times so the results and have to be carefully interpreted.

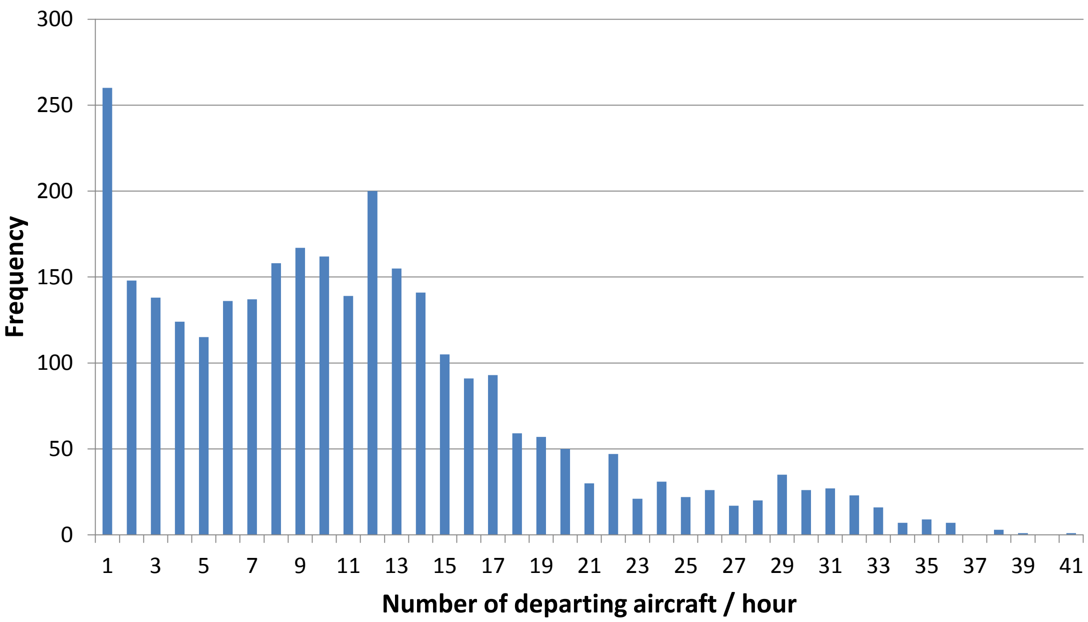


Figure 20: A histogram of the number of departing aircraft / hour during between 1.11.2012 and 31.3.2013

All things considered the presented model has potential in assessing the capacity sufficiency in everyday capacity planning of the deicing process. The model should be further studied for reliability and its forecasting power and the need for further refinement. In the winter 2013-2014, the operating conditions have completely changed as well at the Helsinki Airport so the model isn't applicable to the current operating environment at Helsinki airport.

4.4 Baggage process capacity – Baggage Logistics Centre

Finavia provides the baggage processing service to the airlines. This means moving the baggage from check-in/baggage drop counters to the apron and to the baggage chutes for individual flights. It's the airline's responsibility to move the baggage from the chutes to the aircrafts and load them. In practice, airlines outsource this activity to different airport service providers like:

- Swissport
- Airpro

The moving and loading of the baggage involves manual labor with baggage trucks, containers, trolleys and other equipment. Acquiring new workforce and equipment can increase the capacity for this type of processes, yet this was left out of the study. Instead the baggage process was studied in particular the baggage logistics center which is situated in the terminal 2 part of the baggage processing infrastructure.

The main processes which affect the overall capacity at the *BLC* are identified based on an interview with (Equipment maintenance Manager and Maintenance engineer, 2013) and technical documentation. The processes are check-in and baggage drop-off processing, security screening, baggage sorting. An illustration was created based on the sources, which is shown in Figure 20. The processes are found for both Terminal 1 and Terminal 2 at the airport. A bottleneck could occur in different situations in different parts of the whole baggage processing. In this study the aim is to concentrate on situation when baggage process is near capacity in the current situation.

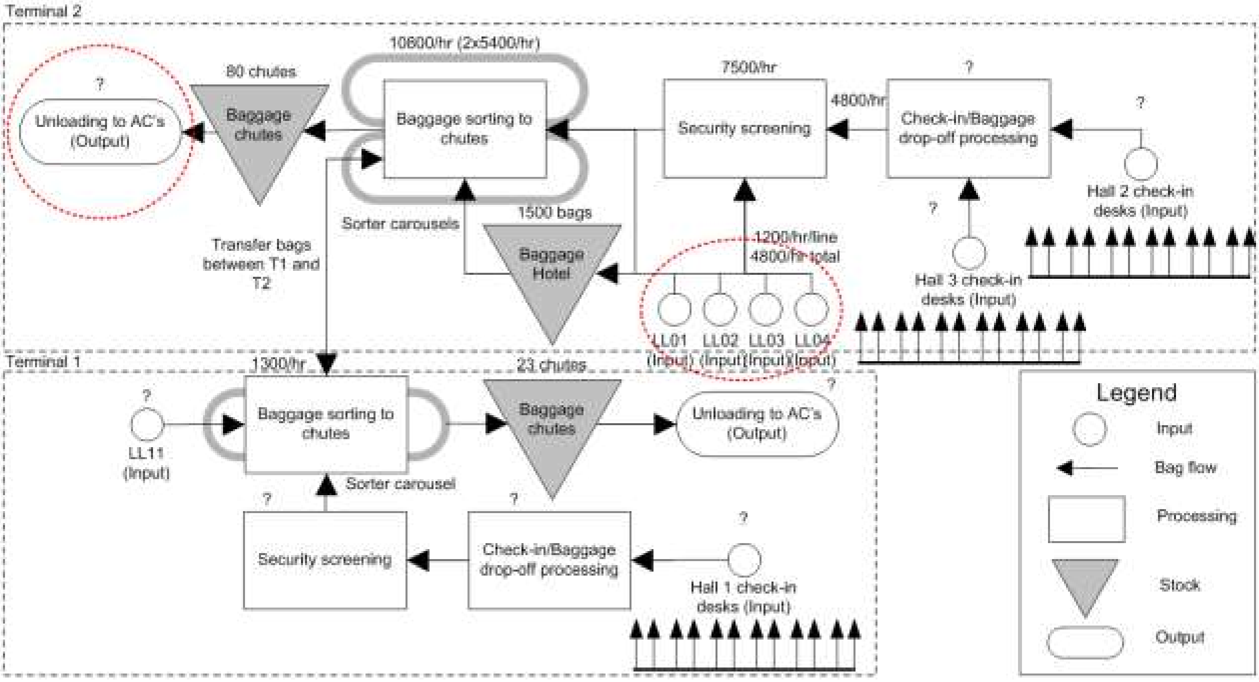


Figure 21: Simplified Process flowchart of the outbound baggage process at Baggage Logistics Center (*BLC*) at Helsinki-Vantaa airport. The identified possible bottlenecks of the process are marked with red circles.

In the interview with Equipment maintenance Manager and Maintenance engineer, 2013 the reasons for which bottlenecks are most likely to occur at the *BLC* were recognized:

- Transfer baggage loading hall capacity at terminal 2 (Loading lines 01-04 in Figure 21). All the long haul flights' baggage are loaded here to the *BLC*
- Baggage unloading hall capacity for departing flights at the terminal 2. The baggage for long haul and non-schengen flights are unloaded from the chutes here.
- Malfunctions and operational breaks in the system. These would include malfunctions in the baggage handling system and information systems related to baggage handling
- A short high peak in the baggage volumes either from the arriving aircrafts or when the apron service providers are unable to balance their capacity with *BLC*:s output

The processes are marked with circles in Figure 21. At the current utilization level, other points in the system are identified of having excess capacity. The reason for the two processes to be identified as bottlenecks was that the physical area at these two points were two small for the baggage trucks and trolleys of the baggage service providers to move and operate at the designed capacity limits. The designed capacities are marked over the respective processes and flows were available or applicable.

In was decided to study further the capacity of the transfer baggage loading hall, because data was available for it and it appeared to be the most restricting factor according to the informants. According to the interview, the practical capacity of the transfer baggage loading hall 500 bags/10 minutes or 3000 bags/hour even though the designed capacity of the hall is 4800 bags/hour.

The data was received from Finavia's baggage process operational database. The data includes the loading lines' loading volume in 10 minute time intervals at the flight peak period for each loading line and each day. Because of difficulties in data retrieval, only the data on previously gathered daily reports was used. The data ranged from 11.02.2013 to 04.07.2013 with total of 188 days of data. The data is then a non-random sample of the loading volumes in the described time period.

The median and different percentiles of the *BLC* transfer baggage loading hall volumes are plotted in Figure 22 as time series. The median of the 10 minute interval volumes is clearly well below the practical capacity of 500/10 min. The time series of the median values fluctuates every day, having maximum value at the time 15:20-15:40. The toughest days are Fridays, Saturdays and Sundays. According to the interview Saturdays have more charter traffic which why there are a lot of bags on Saturdays.

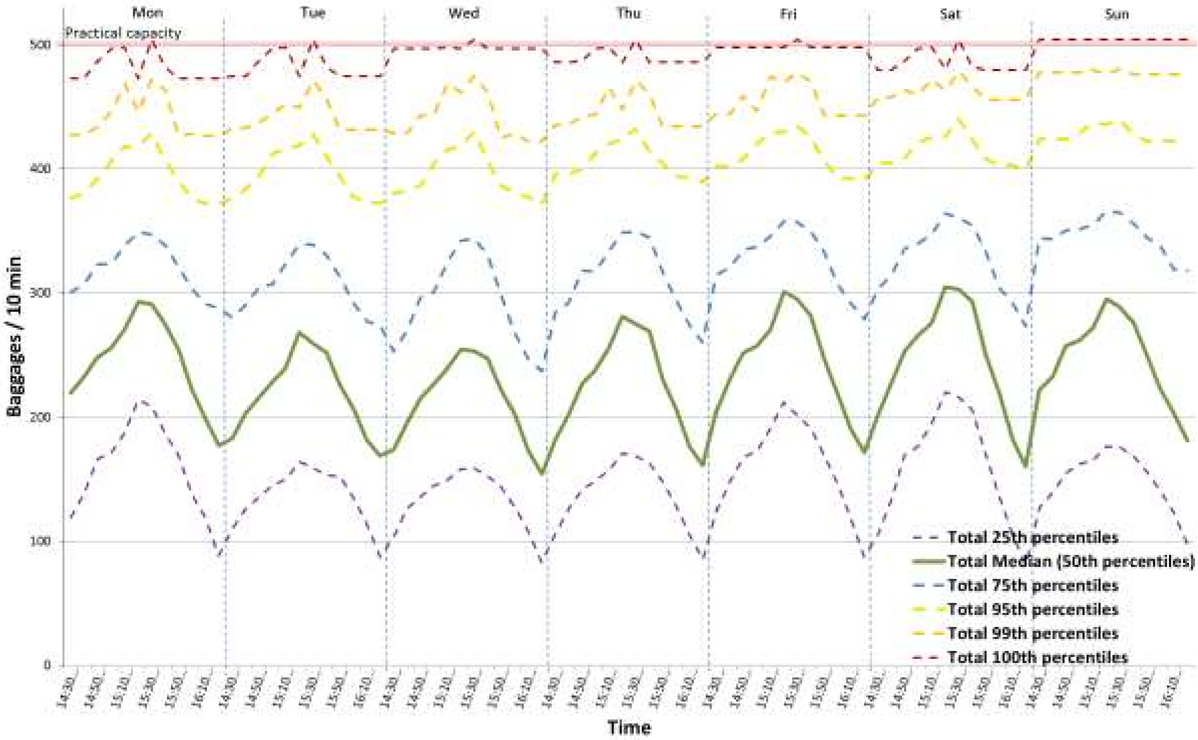


Figure 22: *BLC* transfer baggage loading amount median and percentiles at the demand peak period through week

The 95th percentile is moving around 400 bags/10min line which means that only around 5% of all the peak period times are between 400-500 bags/10 minutes.

The Figure 23 has the loading volumes' medians exploded to the individual loading lines. It can be seen that the loading line 01 is used the most, 02 the second, 03 the third and 04 the least most of the time. The maximum values scattered with red crosses, show on the other hand that the individual loading lines practical capacity is actually much larger than 125 bags/10min which would be expected (500/hr. divided by 4). Clearly even nearly 200 bags per 10 min have been achieved which is the design capacity. This reinforces notion that the physical area around the four loading lines is the reason why the practical capacity is lower than the designed capacity.

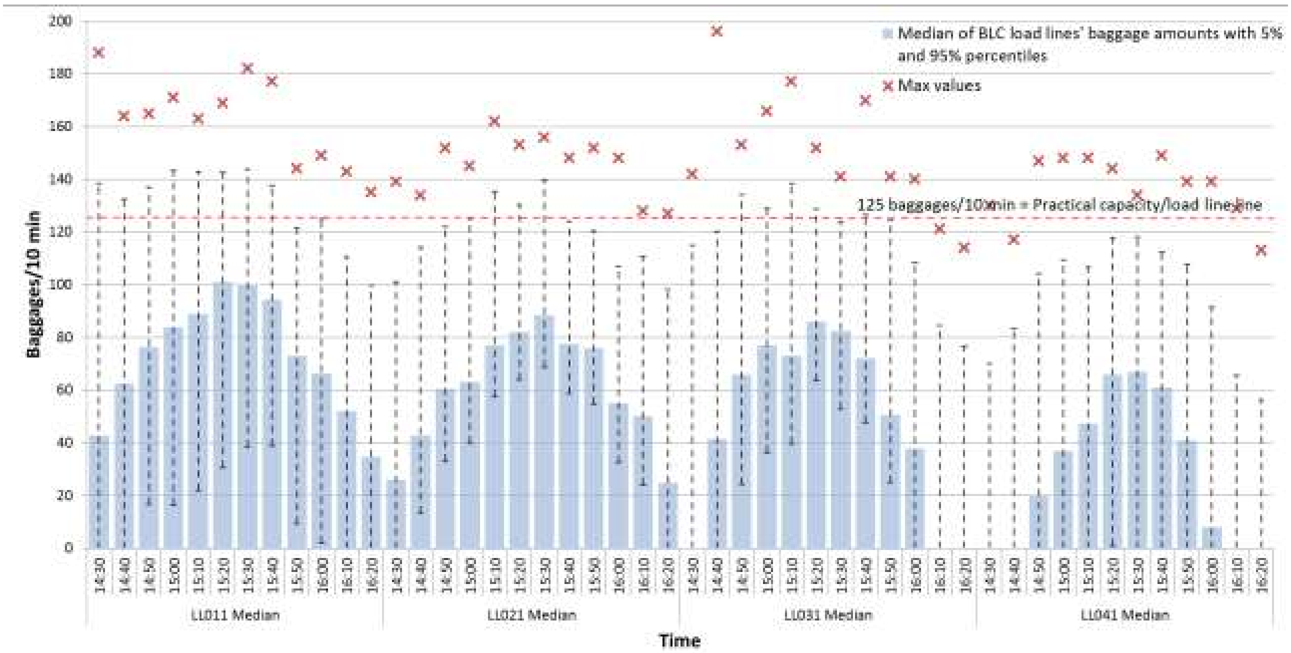


Figure 23: The median, 5% percentile, 95% percentile and maximum loading volumes of BLC transfer loading lines in 10 minute time intervals during the peak period

The volumes in each loading line descend from loading line 11 to 41 in Figure 23: The median, 5% percentile, 95% percentile and maximum loading volumes of BLC transfer loading lines in 10 minute time intervals during the peak period This happens because there is a fifth loading line for special bags beside the first loading line. The operators choose the first available loading line nearest to the first loading line, so they need to walk the shortest distance to load the special bags to the special baggage loading line. The special baggage loading line is then misplaced. In order to minimize to amount needed for the operators to walk to the special loading line it would need to be in the middle of loading hall. The special baggage loading lines place is fixed and cannot be changed without relatively big modification to the building and *BLC* conveyor system.

In Table 10 (*BLC*'s transfer loading lines' capacity utilization measures in current state vs. the estimated maximum utilization) the median, 95th percentile, 99th percentile and maximum values of the individual loading lines' volumes are summarized as well as the total volumes per 10 min time interval. If the baggage volumes on the loading lines were to increase then with the assumptions,

- The increased load is somewhat evenly distributed to the 14:00-16:30 peak period (e.g. the shape of the distribution remains unchanged)

- The variance in the distribution of the 10 min loads will remain same
- The number of workers and baggage towing trucks needed for the amount of bags number are available
- The action plans are modified to support continuous action near capacity.
- The most urgent baggage is prioritized
- All other conditions remain same

it could be expected that the total load volume of bags could be increased by around 100 bags/10 min and still around 95% bags could be delivered without any additional delay or queuing. In practice this would mean that the lines depicting the distribution of the load on capacity in Figure 22 would be transferred upwards by 100 bags/10 min. The estimated expected values are summarized in Table 10: *BLC*'s transfer loading lines' capacity utilization measures in current state vs. the estimated maximum utilization) below title "Estimated maximum capacity with ~95% without queuing".

Table 10: *BLC*'s transfer loading lines' capacity utilization measures in current state vs. the estimated maximum utilization

Measure	Value	95% percentile	99% percentile	Max
Current values				
LL011 bags/10 min median	76	136	161	188
LL021 bags/10 min median	65	124	142,54	162
LL031 bags/10 min median	56	125,5	144	196
LL041 bags/10 min median	56	125,45	144	149
Total bags/10min average	232	412	462	504
Utilization average	46,4 %	82,4 %	92,4 %	101 %
Estimated maximum capacity with ~95% without queuing				
Total bags/10min average	332	512	562	604
Total bags/hour	1991	N/A	N/A	N/A
Peak period total (14:00-16:30)	~5000	N/A	N/A	N/A
Estimated utilization average	66,4 %	102,4 %	112,4 %	120,8 %

Increase in the average of total bag volume from 232 to 332 average total bags /10min would mean average loading volume increase from around 1400 bags/hour to around 2000 bags/hour. For the whole peak period (14:00-16:30 in baggage handling) the increase would be then from 3500 bags on the peak period to 5000 bags. When presented in fractional terms, a 43% increase in average volume would possible according to this analysis which would mean around five fully loaded wide body aircrafts arriving and departing during the peak hour.

When operating at increased level and around five percent of the bags coming to the loading line hall would experience queuing. In theory the distribution of the total loading lines volumes would implicate that the amount of bags at the same time coming to the transfer bag loading hall would be from 0-100. In practice the amount would be the amount that can be fitted to a baggage truck and trolleys. Throughput time for the 5% of the bags (from 0 to 150 bags) would increase for around 0 to 10min, which is the time taken to empty one baggage delivery.

4.4.1 Error sources

The actual input volume doesn't necessarily reflect the demand at the *BLC* Transfer baggage loading area. The demand means in this case the baggage towing trucks driving to the transfer bags loading area. There is no data available on the arriving trucks. For example, it is not certain in the light of the data has there been any queuing at the facility. If the loading area is already full and at near capacity volumes and the arriving baggage towing truck has to wait outside to get in, because of the physical limitations at the loading area.

5. Conclusions and discussion

The results and conclusions based on the findings are presented here.

The study was made based on the prevailing operating conditions during the information gathering period of the study. The effects of operational conditions changes are left for further study and it might be easier to conduct based on the results of this thesis. The conclusions and discussions of the findings are presented under their respective titles.

The studying of the capacities at an airport proved to be challenging task. The problem arises as there are many different states of the system and so many different situations can occur in the complex system. An airport's operations are variable and experience a lot of randomness as well. The capacity can be affected also by events taken place long time ago. This is why it is hard to tell what amount of capacity is sufficient as exceptional situations rise where there are suddenly need for huge amounts of capacity.

The choosing of the models, data gathering and data processing and combining took longer than expected. The need to create custom plots with R-software to present the results was also challenging and time consuming. The scope of the study was too broad to drill down in detail to all the studied processes.

5.1 Helsinki Airport process model

The introduced process model of the Helsinki Airport depicts linkages between the main flows of objects (excluding the freight) from one process to another, as well as the primary processes doing processing on the flows. The Helsinki Airport is depicted as a system that has inputs and outputs of the main flows of object and in the system the flows either converge or diverge. The capacity needs to be available from one processing stage and movement to the next if there is no spare time available. The diagram is able to communicate the linkages between consecutive processes and how they are dependent from each other.

There is great complexity in the Helsinki Airport as system. The complexity makes it hard to assess the sufficiency of capacities of consecutive processes and movements from one state or process to another. The capacities vary also within the high level processes presented as there are multiple activities that can be seen as having their own capacities. One influencing factor in assessing capacities is the service supply network of the airport. The inter-organizational linkages have their own impact on the capacity considerations. Services bought by the airlines are provided by different service suppliers. For example the capacities of the turnaround operations are different for the airline companies depending on their service supplier's capacity.

The model process diagram could be used e.g. as basis for creating a discrete-event simulation model of the whole operation of Helsinki Airport. Discrete-event simulation models are largely used for studying the capacities and planning for capacity at airports (Ashford et al., 2011, p.602). Simulation models are the only way to incorporate the whole airport system in to a model and study the effects of individual parts jointly on the system as whole. Today no airport development is done without using airport modeling and simulation (Ashford et al., 2011, p. 602).

The created process model of Helsinki Airport needs to be updated if changes in processes or infrastructure happen at the airport. The model depicts the state of the processes as they were during the data collection stage of the study.

5.2 Runway process capacity

Modeling the runway capacity with piecewise linear quantile regression seems to give good and pretty reliable results. The different levels of the factors' effects on the capacity modeled result in capacity curves that don't overlap each other and the shapes indicate that possibility. The modeling method seems to be good for the estimation of the runway capacities in different conditions.

Biggest capacities are achieved with runway configurations with the configuration 22L/22R. The next ones are 04L/04R, 04R/04R and 15/22R. As the runway configuration is selected by the Air traffic control on operational grounds, the configuration with best capacity cannot be selected in all operating conditions to maximize the capacity for conducting all the scheduled runway operations. Instead the information on the capacities can be used to evaluate the effect of the runway configuration in use to the daily flight schedules at Helsinki Airport. The information of runway configuration capacities can also be used to isolate the effects of other factors, as was done in this study with the effects of snowfall, wind speed and visibility on the runway capacity.

The results indicate that the planned runway usage exceeded in many occasions even the theoretical capacities that this study suggests. Some of the planned runway operations seem not have been possible to be conducted in the given time frame even in the best possible capacity conditions. This would indicate that some of these scheduled runway operations have had to have been operated at an earlier or later stage in time than when they were scheduled. Because some might have taken place earlier and some later than scheduled, the result don't directly implicate that delays of flights have incurred due to this, but that there is a possibility of occurrence of delays because of over planned capacity utilization. The study of this possibility is left for further research.

The effect of snowfall on the runway capacity was evaluated and it seems to have a great impact on the capacity of the runways. Large amounts of snowfall create potential for a bottleneck in the runway operations if incurred during the peak operating times at Helsinki Airport. The capacity decreases rapidly as the snowfall increases and the effects of snowfall on the capacity range from 1 to 8 reduced operations per 15 minutes in different situations

The effect of high average wind speed on the runway capacity seems to be quite high. Bottlenecks might occur due to the high wind speeds at the peak operating times, but the results don't seem to be that reliable and because of that the effects could be studied further.

The effect of visibility has only modest effect on the capacity reduction. In most cases it is not a potential factor to create a bottleneck in the runway operations, as the effect size is small (around 1 operation reduction in most cases) and as the highest amounts of runway operations are relatively quite rare (there are typically only few occasions on a day) when the low visibility could be constraining the number of runway operations to be carried out.

It is important to note that the studied effects might happen simultaneously and their joint effects were not studied. There are also many error sources in the estimation, which were covered in the findings section.

5.3 Deicing process capacity

The deicing capacity was estimated inversely by creating a model for the expected deicing treatment time. A gamma regression model was used for the expected treatment time estimation and the different factors effects on the processing time. From the treatment time model an equation for calculating the expected capacity in different situations was derived. An example calculation with the derived equation for capacity calculation was presented and the results compared with the demand for departures data date interval. The model seems to be pretty good in terms of statistical inference and regression diagnostics, but not a perfect fit.

The effects of deicing location, aircraft type, number of deicing trucks, snowfall, temperature, average wind and relative humidity on the expected deicing time were studied and estimated. The factors effects on the capacity were compared with the demand for departing flights in the data time range and the capacity sufficiency in each scenario assessed.

The results indicate that deicing process could become constraining in multiple of situations. The model would indicate that theoretically around 10 treatment trucks would be sufficient to cover the need for capacity in almost all of the demand situations for the departing aircraft. However it has to be kept in mind that the model doesn't include the changing times (the time between consecutive aircraft deicing treatments) so the capacity estimates are probably too optimistic and therefore only theoretical.

5.4 Baggage process capacity

The capacity of the baggage center was estimated for the most potential bottleneck point, the transfer baggage loading line of the *BLC*. Other identified potential bottleneck point in the *BLC* was the baggage unloading area, but capacity estimated or data was unavailable. The study of capacity of the complete baggage facility was decided to be left out of this study.

The baggage unloaded from the aircraft to the terminal 2, are loaded into the baggage system through transfer baggage loading lines. The capacity was identified in the discussions with the *BLC* staff as 500 baggage/10 minutes for the standard baggage. The data from the *BLC* information systems supported this estimate. In addition it was estimated that on average the utilization of this possible bottleneck point could probably be increased at most by around 100 baggage/10 minutes. This figure would suggest that the baggage system at the state during study could at most take 5 fully loaded wide body aircraft arriving and departing during the peak period before the capacity limits of the loading lines would significantly constrain the operation. The most important assumptions made to come to this conclusion are that the arrivals of these aircraft would be evenly distributed on the peak period and that the baggage unloading area capacity would suffice as well.

It was also identified that the transfer baggage loading lines' capacity was limited by the physical size of the hall around the loading lines. By increasing the physical area the maximum capacity and creating operational changes could probably be further increased to near the design capacity of 800 baggage/15 minutes. On the other hand it might not be possible to increase the size of the loading, because the area is situated at a place in the terminal building where there is little room to expand.

5.5 Further study

In order to fully assess the reliability of the findings of this study it would be fruitful that the results on the studied processes were validated by other experts. The runway capacity analyses could be verified; e.g. the reasons behind the dropping of capacity in the conditions presented could be studied to better understand the phenomena.

Also deicing process model requires further validation and the forecasting power of model further testing in order to be taken into wider use. This type of an inverse model for estimating capacity was used, as the problems with missing data prevented the direct estimation of the capacity (e.g. deicing treatments/hour).

The including of the change times in the deicing processing could be studied further and the calculation of expected capacity developed further. Especially problematic for the expected capacity calculation model is that it is not including the change times between consecutive aircraft to be treated by a certain processing line. The model could also be fitted and applied to the current operating conditions in the deicing process which have changed since the data was gathered.

In the baggage process it could be studied what is the capacity of the baggage unloading facility. It was also identified as a possible bottleneck point under certain circumstances. However, it was left out of the study because of the difficulty of acquiring data on the unloading activities. Also more thorough research on the capacity could be conducted on the whole baggage logistics centre.

The baggage process capacity could as well be more broadly studied. Questions like what is the acceptable level of throughput time delay for baggage or what is the throughput time in different scenarios, and how they affect the delays of i.e. departing flights should be answered.

Also the other processes at Helsinki Airport could be studied and modeled, perhaps simulated. A whole model of Helsinki Airport could be developed with discrete-event simulation tools created specifically for airport environment. This type of modeling is very laborious and requires a lot of resources, but could provide more thorough understanding of the complex system of a complete airport.

6. References

Adler, D., Murdoch, D., 2013. rgl: 3D visualization device system (OpenGL).
Anderson, V., Johnson, L., 1997. Systems thinking basics: from concepts to causal loops. Pegasus Communications, Cambridge, Mass.
Ashford, N.J., Mumayiz, S., Wright, P.H., 2011. Airport Engineering: Planning, Design and Development of 21st Century Airports. John Wiley & Sons.
Bozarth, C., Handfield, R.B., 2005. Introduction to Operations and Supply Chain Management, 1st ed. ed. Prentice Hall.
Doganis, R., 2001. The airline business in the twenty-first century. Routledge, London; New York.
Duty Manager, HUB control center, 2013. Interview about Helsinki Airport, Finnair Oyj.
Equipment maintenance Manager, Maintenance engineer, 2013. Interview about Baggage Logistic Center capacity,Finavia Oyj.
Finavia Oyj, 2013a. Traffic statistics [WWW Document]. URL http://www.finavia.fi/en/news-room/statistics/ (accessed 1.31.14).
Finavia Oyj, 2013b. Aeronautical Information Services [WWW Document]. URL https://ais.fi/ais/eaip/html/efhk.htm (accessed 10.14.13).
Finnair Oyj, 2013. Annual Report 2012 [WWW Document]. URL http://www.finnairgroup.com/konserni/konserni_12_2.html (accessed 2.3.14).
Finnish Meteorological Institute, 2013. Data catalog of weather observations at Helsinki-Vantaa airport [WWW Document]. URL http://catalog.fmi.fi/geonetwork/srv/en/main.home (accessed 10.1.13).
Fox, J., 2008. Applied Regression Analysis and Generalized Linear Models. SAGE Publications.
Gilbo, E.P., 1993. Airport Capacity: Representation, Estimation, Optimization. IEEE Trans. Control Syst. Technol. 1.
Goldratt, E.M., Cox, J., 1993. The Goal: A Process of Ongoing Improvement. Gower.
Johnston, R., Clark, C., 2008. Service operations management: improving service delivery, 3rd ed. Pearson Education.
Koenker, R., 2005. Quantile Regression, Econometric Society Monographs. Cambridge University Press.
Koenker, R., 2013. Quantile regression in R: A Vignette.
Koenker, R., Hallock, K.F., 2001. Quantile Regression. J. Econ. Perspect. 15, 143–156.
Koenker, R., Portnoy, S., Tian Ng, P., Zeileis, A., Grosjean, P., Ripley, B.D., 2013. quantreg: Quantile Regression.
Krajewski, L., 2007. Operations Management: Processes And Value Chains, 8th ed. Pearson Education.
Lambert, D.M., Cooper, M.C., 2000. Issues in Supply Chain Management. Ind. Mark. Manag. 29, 65–83. doi:10.1016/S0019-8501(99)00113-3
Myers, R.H., 2010. Generalized linear models: with applications in engineering and the sciences. Wiley, Hoboken, N.J.

Ramanujam, V., Balakrishnan, H., 2009. Estimation of arrival-departure capacity tradeoffs in multi-airport systems. IEEE, pp. 2534–2540. doi:10.1109/CDC.2009.5400462
Siha, S., 1999. A classified model for applying the theory of constraints to service organizations. Manag. Serv. Qual. 9, 255–264. doi:10.1108/09604529910273201
Watson, K.J., Blackstone, J.H., Gardiner, S.C., 2007. The evolution of a management philosophy: The theory of constraints. J. Oper. Manag. 25, 387-402. doi:10.1016/j.jom.2006.04.004

Appendix A: R – function: hist3d() for 3D-histogram plotting

```
hist3d<-function(x,y,nclass="auto",alpha=1,tcol="#ff0000", scol=
"#aaaaaa", scale=10, csize=0,xlims=c(0,max(x,y)),zlims=c(0,max(x,y)))
{
    save <- par3d(skipRedraw=TRUE)
    rgl.bg(col="transparent")
    xy <- xy.coords(x,y)
    x <- xy$x
    y <- xy$y
    n<-length(x)

    breaks.x <- seq(min(x)-0.5,16 + 0.5,length=(16-min(x)+2))
    breaks.y <- seq(min(y)-0.5,16 + 0.5,length=(16-min(y)+2))

    z<-matrix(0,(16-min(x)+1),(16-min(y)+1))
    for (i in 1:(16-min(x)+1))
    {
    for (j in 1:(16-min(y)+1))
      {
      z[i, j] <- sum(x < breaks.x[i+1] & y < breaks.y[j+1] &
                      x >= breaks.x[i] & y >= breaks.y[j])

binplot.3d(c(breaks.x[i],breaks.x[i+1]),c(breaks.y[j],breaks.y[j+1]),
                 z[i,j],alpha=alpha,topcol=tcol, sidecol=scol,
           size=csize, xlim=16, zlim=16, ylim=110) #xlim=xlims,
           zlim=zlims)
      }
    }
    aspect3d(1, 0.5, 16/16)
    xlength=max(x)-min(x)+1
    ylength=max(y)-min(y)+1
              bbox3d(xlim=xlims, zlim=zlims,
           ylims=c(0:110),xat=c(0:15),zat=c(0:15),
           yat=c(seq(0,100,20)),front="lines", back="lines", col-
           or="black", lit=FALSE, expand=1.05, cex=1.2, marklen=30)

    grid3d('y-', at=list(x=(0:16), z=(0:16)))
    grid3d('x')
    grid3d('z')
    rgl.viewpoint(theta=22.5,phi=30, fov=20,zoom=1)
    light3d(theta = 60, phi = 45)
    par3d( windowRect=c(20,20, 720,720),skipRedraw=FALSE)
    return(list(freqs=z))
}
```

Appendix B: R – function: binplot.3d() for 3D-histogram plotting

```
binplot.3d<-
func-
tion(x,y,z,alpha=1,topcol="#ff0000",sidecol="#aaaaaa",size=0,xlim=c(0,
max(x,y)),zlim=c(0,max(x,y)), ylim=c(0, 110))
{
  save <- par3d(skipRedraw=TRUE)
  on.exit(par3d(save))

  x1<-c(rep(c(x[1]+size,x[2]-size,x[2]-
size,x[1]+size),2),rep(x[1]+size,4),rep(x[2]-size,4))
  z1<-c(rep(c(0,0,z,z),4))
  y1<-c(rep(y[1]+size,4),rep(y[2]-size,4),rep(c(y[1]+size,y[2]-
size,y[2]-size,y[1]+size),2))
  x2<-c(rep(c(x[1]+size,x[1]+size,x[2]-size,x[2]-
size),2),rep(c(x[1]+size,x[2]-size,rep(x[1]+size,3),rep(x[2]-
size,3)),2))
  z2<-c(rep(c(0,z),4),rep(0,8),rep(z,8) )
  y2<-c(rep(y[1]+size,4),rep(y[2]-
size,4),rep(c(rep(y[1]+size,3),rep(y[2]-size,3),y[1]+size,y[2]-
size),2) )
  rgl.quads(x1,z1,y1,col=rep(sidecol,each=4),alpha=alpha, xlim=xlim,
zlim=zlim, ylim=ylim)
  if( !(z==0) )
  {
    rgl.quads(c(x[1]+size,x[2]-size,x[2]-
size,x[1]+size),rep(z,4),c(y[1]+size,y[1]+size,y[2]-size,y[2]-size),
    col=rep(topcol,each=4),alpha=1,xlim=xlim, zlim=zlim, ylim=ylim)
    rgl.lines(x2,z2,y2,col="#000000", lwd=1, xlim=xlim, zlim=zlim,
ylim=ylim)
  }
}
```

Appendix C: An example R – Script for fitting and plotting runway capacity curves

```
setwd( "C:/Users/okaplas/Desktop/Plots")
env = globalenv()
fittedQR=list()

colors= c( "#1B9E77" , "#7570B3" ,"#D95F02" )
noColors=length(colors)


legends=NULL

quantile=0.999

peak=FALSE
peakStart="12:00"
peakEnd="15:30"
atLeastOnePlot=FALSE

comb=c(  "04L/04R", "04R/04R",  "15/22R","22L/22R")
# All: c( "04L/04L", "04L/04R", "04L/15",  "04L/22R", "04R/04R",
"04R/22R", "15/04R",  "15/15",   "15/22L",  "15/22R",  "22L/04R",
"22L/15",  "22L/22L","22L/22R", "33/22R" , "33/33")
# Data available: c( "04L/04L", "04L/04R", "04L/15",  "04L/22R",
"04R/04R", "04R/22R", "15/04R",  "15/15",   "15/22L",  "15/22R",
"22L/04R", "22L/15",  "22L/22L","22L/22R", "33/22R" , "33/33")


snowLevel=levels(runwayStatistics[,9])
snowLevel<-c(snowLevel[1],snowLevel[3],snowLevel[4],snowLevel[2])
visibilityLevel=levels(runwayStatistics[,12])
windLevel=levels(runwayStatistics[,11])

plots<-NULL
plots<-list()
#runwayData<-list()

#ggdf<-list()
ggdf<-NULL

statisticsRunwayCombWeather<-NULL
statisticsRunwayCombWeather<- data.frame(RunwayComb=character(),
SnowLevel=character(),WithinCapacity=numeric(),OutsideCapacity=numeric
(), WithinPercentOfTotal=numeric(), stringsAsFactors=FALSE)

for ( i in 1: length(comb))
{

    atLeastOnePlot=FALSE



    for(w in 1:length(snowLevel))
    {
```

```
        weather=snowLevel[w]
        if(peak)
            runwayDa-
ta=runwayStatistics[runwayStatistics["RwyCombHour"]==comb[i] & run-
wayStatistics["snowCategory"]==weather & format(runwayStatistics[,1],
"%H:%M")>= peakStart & format(runwayStatistics[,1], "%H:%M")<= peakEnd
,]
        else
            runwayDa-
ta=runwayStatistics[runwayStatistics["RwyCombHour"]==comb[i] & run-
wayStatistics["snowCategory"]==weather ,]

        if(weather == "")   weather="No snow"

        if(nrow(runwayData) < 20 | weather == "melt")
        {
            if(is.null(legends))
            {
                assign(paste("legends_",i,sep=""), c(), envir=env)
                assign(paste("legendColors_",i,sep=""), c(), en-
vir=env)
            }

            next
        }

        maxArr<-max(runwayData["ARR"])
        maxDep<-max(runwayData["DEP"])
        open3d(family="mono",cex=1, antialias=4)

        fit<-rqss( DEP ~ qss( ARR , lambda = 0.01,constraint="CD"),
tau = quantile, data=runwayData)
        fittedQR[[i]]<-fit

        histogram<-hist3d(runwayData[,13],runwayData[,14], alpha=1,
nclass=17, tcol='lightblue1', scol='lightblue1',csize=0.4,
xlims=c(0,15),zlims=c(0,15))
        maxFreq=max(histogram$freqs)

        depValues=c(fit$coef[1] , fit$coef[1] + fit$coef[-1] )
        dif = maxArr + 1 - length(depValues)



        if(exists(paste("legends_",i, sep="")))
        {
            assign(paste("legends_",i,
sep=""),c(get(paste("legends_",i,sep=""), envir=env),
paste(comb[i],weather,"N=",nrow(runwayData))), envir=env)
            assign(paste("legendColors_",i,
sep=""),c(get(paste("legendColors_",i, sep=""), envir=env), col-
ors[w%%noColors + 1]), envir=env)
        }
        else
        {
            assign(paste("legends_",i, sep=""),c(
paste(comb[i],weather,"N=",nrow(runwayData))), envir=env)
            assign(paste("legendColors_",i, sep=""),c( col-
ors[w%%noColors + 1]), envir=env)
```

```
	}

	if(!is.null(ggdf))
numOfCurves=length(unique(ggdf[ggdf[,"RunwayComb"] == comb[i], "group"
]))


	capacityCurve<-function(arr, DEP)
	{
	    if(length(DEP) - 1 < arr) return(0)

	    decimals<-arr-floor(arr)
	    return ( (1-
decimals)*DEP[floor(arr)]+decimals*DEP[round(arr)])
	}

	isWithinCapacity<-function(arr, dep, DEP)
	{
	    if(length(DEP) - 1 < arr){ return(FALSE) }

	    decimals<-arr-floor(arr)

	    if( (1-
decimals)*DEP[floor(arr)+1]+decimals*DEP[round(arr)+1] >= dep) re-
turn(TRUE)
	    else return(FALSE)
	}

	demand<-runwayPlannedLoadPeak[ run-
wayPlannedLoadPeak$RunwayComb==comb[i],]

	withinCapacityCumul=0
	outsideCapacityCumul=0
	DEP<-c(fit$coef[1] ,fit$coef[1]+fit$coef[-1])
	for( a in 1:nrow(demand) )
	{
	    if(isWithinCapacity(demand[a,"ARR"], demand[a,"DEP"],
DEP))
	        withinCapacityCu-
mul=withinCapacityCumul+demand[a,"Frequency"]
	    else
	        outsideCapacityCu-
mul=outsideCapacityCumul+demand[a,"Frequency"]
	}
	print("Comb   snowLevel   withinCapacity  outsideCapacity
within%")
	print(paste(comb[i],snowLevel[j],withinCapacityCumul, outside-
CapacityCumul, withinCapacityCu-
mul/(withinCapacityCumul+outsideCapacityCumul), sep=";"))

	if(nrow(statisticsRunwayCombWeather) == 0 )
	    statisticsRunwayCombWeather<-data.frame( Runway-
way-
Comb=comb[i],SnowLevel=snowLevel[w],WithinCapacity=withinCapacityCumul
, OutsideCapacity=outsideCapacityCumul, WithinPercentOfTo-
tal=withinCapacityCumul/(withinCapacityCumul + outsideCapacityCu-
mul)*100, stringsAsFactors=FALSE)
	else
```

```
            statisticsRunwayCombWeather<-
rbind(statisticsRunwayCombWeather, list( Runway-
way-
Comb=comb[i],SnowLevel=snowLevel[w],WithinCapacity=withinCapacityCumul
, OutsideCapacity=outsideCapacityCumul,  WithinPercentOfTo-
tal=withinCapacityCumul/(withinCapacityCumul+outsideCapacityCumul)*100
))
        print(paste(comb[i],visibilityLevel[w],withinCapacityCumul,
outsideCapacityCumul, withinCapacityCu-
mul/(withinCapacityCumul+outsideCapacityCumul), sep=";"))

        if( dif > 0 )
        {
          lines3d(c(0:maxArr), rep(0,maxArr+1),c(fit$coef[1]
,fit$coef[1]+fit$coef[-1], rep(0, dif)), color="red",
lwd=2,xlims=c(0,16),zlims=c(0,16))

          if( is.null(ggdf))
          {
            ggdf<-data.frame(ARR=c(0:maxArr), DEP=c(fit$coef[1]
,fit$coef[1]+fit$coef[-1], rep(0, dif)), Run-
wayComb=rep(comb[i],maxArr+1) , col=colors[w],
group=weather,legend=rep( paste("n=",nrow(runwayData),sep=""), max-
Arr+1), legendYPos=rep(0.9*16, maxArr+1) )
          } else if (numOfCurves < 1)
          {
            ggdf<-rbind(ggdf,data.frame(ARR=c(0:maxArr),
DEP=c(fit$coef[1] ,fit$coef[1]+fit$coef[-1], rep(0, dif)), Run-
wayComb=rep(comb[i],maxArr+1), col=colors[w],
group=weather,legend=rep( paste("n=",nrow(runwayData),sep=""), max-
Arr+1), legendYPos=rep(0.9*16, maxArr+1) ))
          } else
          {
            ggdf<-rbind(ggdf,data.frame(ARR=c(0:maxArr),
DEP=c(fit$coef[1] ,fit$coef[1]+fit$coef[-1], rep(0, dif)), Run-
wayComb=rep(comb[i],maxArr+1), col=colors[w],
group=weather,legend=rep( paste("n=",nrow(runwayData),sep=""), max-
Arr+1), legendYPos=rep((0.9-numOfCurves*0.07)*16, maxArr+1) ))
          }

        }
        else
        {
          lines3d(c(0:maxArr, maxArr+0.01),
rep(0,maxArr+2),c(fit$coef[1] ,fit$coef[1]+fit$coef[-1], 0), col-
or="red", lwd=2,xlims=c(0,16), zlims=c(0,16))

          if( is.null(ggdf))
          {
              ggdf<-data.frame(ARR=c(0:maxArr, maxArr+0.01),
DEP=c(fit$coef[1] ,fit$coef[1]+fit$coef[-1],0), Runway-
way-
Comb=rep(comb[i],maxArr+2),col=colors[w],group=weather,legend=rep(past
e("n=",nrow(runwayData),sep=""), maxArr+2),legendYPos=rep(0.9*16, max-
Arr+2))
          }
          else if (numOfCurves < 1)
          {
```

```
            ggdf<-rbind(ggdf,data.frame(ARR=c(0:maxArr, max-
Arr+0.01), DEP=c(fit$coef[1] ,fit$coef[1]+fit$coef[-1],0), Run-
wayComb=rep(comb[i],maxArr+2),col=colors[w],group=weather, leg-
end=rep(paste("n=",nrow(runwayData),sep=""), max-
Arr+2),legendYPos=rep(0.9*16, maxArr+2) ))
          }
          else
          {
            ggdf<-rbind(ggdf,data.frame(ARR=c(0:maxArr, max-
Arr+0.01), DEP=c(fit$coef[1] ,fit$coef[1]+fit$coef[-1],0), Run-
wayComb=rep(comb[i],maxArr+2),col=colors[w],group=weather, leg-
end=rep(paste("n=",nrow(runwayData),sep=""), max-
Arr+2),legendYPos=rep((0.9-numOfCurves*0.07)*16, maxArr+2) ))
          }
        }
        atLeastOnePlot=TRUE

        grid3d('y-', at=list(x=(0:16), z=(0:16)))
        grid3d('x')
        grid3d('z')

        mtext3d(paste("Runway combination ARR" ,sub("/","/DEP
",comb[i])), edge='Y--', at=maxFreq*10/6, color='black', font=2, fami-
ly="mono")
        if(weather != "")
            mtext3d(paste("N=",nrow(runwayData),"Weather :" ,weather),
edge='Y--', at=maxFreq*8/6, color='black', font=2, family="mono")

        mtext3d(paste("Arrivals/15
min,",sub("/.*","",comb[i])),edge='X-+', line=4 , color='black', fami-
ly="mono")
        mtext3d(paste("Departures/15
min,",sub(".*/","",comb[i])),edge='Z+-', line=4 , color='black', fami-
ly="mono" )
        rgl.postscript(paste("RunwayCombHistogram_", gsub(
"/","_",comb[i]) ,"_", gsub( " " ,"",weather),".pdf", separator="",
collapse=""), fmt="pdf")
        rgl.snapshot(paste("RunwayCombHistogram_", gsub(
"/","_",comb[i]) ,"_", gsub( " " ,"",weather),".png", separator="",
collapse=""), fmt="png")


    }
    if(!atLeastOnePlot)
    {
        #dev.off()
        next
    }
}



ggdf[,"Type"]<-"Capacity"
ggdf[, setdiff(names(runwayPlannedLoadPeak), names(ggdf))]<-NA
ggdf<-rbind(ggdf,runwayPlannedLoadPeak)

pdf("RunwayCapacity_EffectOfSnow2.pdf", pointsize=1, width=7.5,
height=7.5)
```

```
gridPlot<-ggplot(subset(ggdf, Type %in% c( "Capacity" ) ),aes(
x=ARR,y=DEP, ymin=0, ymax=DEP,fill=group, group=group,colour=group))
gridPlot<-(gridPlot + geom_line(size=0.5) + geom_ribbon(alpha=0.2) +
xlab(expression("Arrivals/15 min, C"["a"])) +
ylab(expression(paste("Departures/15 min, ", phi,"(",C[a],")"))) +
ggtitle("Snow - Effect on runway capacity")
    + theme(legend.title=element_blank(),legend.text = element_text(
size = 10), legend.position="top",
        plot.background=element_rect(fill="transparent",colour=NA),
pan-
el.background=element_rect(fill=rgb(col2rgb("lightgray")[1],col2rgb("l
ightgray")[2],col2rgb("lightgray")[3], 80, maxColorValue=255)), #pan-
el.background=element_rect(fill="transparent", colour=NA),
        panel.grid.major=element_line(color="white", size=1, line-
type=1) , panel.grid.minor=element_line(color="white", size=1, line-
type=1), #element_line(colour="lightgrey", size=0.5))
        axis.text=element_text(colour="black", size=8), ax-
is.ticks=element_line(colour="black"), ax-
is.title=element_text(size=10),
        plot.title=element_text(size=12))
    + scale_y_continuous(expand = c(0,0),limit=c(0,16), mi-
nor_breaks=c(seq(0,17,1)), breaks=c(seq(0,17,5)) )
    + scale_x_continuous(expand = c(0,0),limit=c(0,16), mi-
nor_breaks=c(seq(0,17,1)), breaks=c(seq(0,17,5)) )
    + scale_colour_manual(values =colors)#, la-
bels=get(paste("legends_",i,sep=""),envir=env))
    + scale_fill_manual(values=colors)#, la-
bels=get(paste("legends_",i,sep=""),envir=env))
    + coord_fixed()
    + facet_wrap(~RunwayComb, ncol=2)
    + geom_rect(aes(xmin=0.68*16 , xmax=0.95*16, ymin=legendYPos-
0.04*16, ymax=0.94*16), color="white", fill="white", linetype=0)
#ymin=legendYPos-0.05*16
    + geom_text(aes(x = 0.7*16, y = legendYPos, label = legend, col-
or=group, size=5, hjust=0), show_guide=FALSE, size=4) #family = "ser-
if"

    + geom_point(data=subset(ggdf, Type %in% c( "Planned" )), aes(
x=ARR, y=DEP , size=sqrt((Frequency)/pi), fill=NA), col-
or=rgb(col2rgb("Black")[1],col2rgb("Black")[2],col2rgb("Black")[3],
150, maxColorValue=255),shape=21, show_guide=FALSE )
    + scale_size_continuous(range=c(1,15))
    + facet_wrap(~RunwayComb, ncol=2)
    )
gridPlot
dev.off()

print("Snow level; Frequency; Total N;Relative frequency;")
relativeFrequencies<-"Snow level; Frequency; Total N;Relative frequen-
cy;\n"
for(b in 1:length(snowLevel))
{

n=nrow(runwayStatistics[runwayStatistics["snowCategory"]==snowLevel[b]
,])
    relativeFrequencies<-paste(relativeFrequencies , snowLevel[b], n,
"9680", n/9680, "\n",sep=";")

}
```

```
cat(relativeFrequencies)
print(statisticsRunwayCombWeather)
```